PF-0459 US
09/002,485
Reference No. 17

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07K 14/00

A2

(11) International Publication Number: WO 98/42739

(43) International Publication Date: 1 October 1998 (01.10.98)

(21) International Application Number: PCT/US98/05653

(22) International Filing Date: 20 March 1998 (20.03.98)

(30) Priority Data:

08/822,167	21 March 1997 (21.03.97)	US
09/044,466	19 March 1998 (19.03.98)	US

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

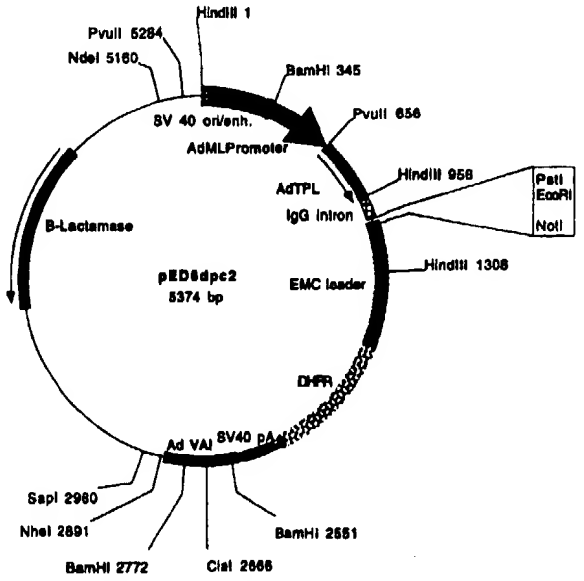
(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 113 Ann Lee Road, Harvard, MA 01451 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Novel polynucleotides and the proteins encoded thereby are disclosed.



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

This application is a continuation-in-part of application Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/822,167), filed March 21, 1997, which is incorporated by reference herein.

FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 707 to nucleotide 1783;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 368 to nucleotide 838;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bp783_3 deposited under accession number ATCC 98369;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bp783_3 deposited under accession number ATCC 98369;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bp783_3 deposited under accession number ATCC 98369;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bp783_3 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 174 to amino acid 183 of SEQ ID NO:2;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 707 to nucleotide 1783; the nucleotide sequence of SEQ ID NO:1 from nucleotide 368 to nucleotide 838; the nucleotide sequence of the full-length protein

coding sequence of clone bp783_3 deposited under accession number ATCC 98369; or the nucleotide sequence of a mature protein coding sequence of clone bp783_3 deposited under accession number ATCC 98369. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone bp783_3 deposited under accession number ATCC 98369. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 44.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 44;

(c) fragments of the amino acid sequence of SEQ ID NO:2 comprising the amino acid sequence from amino acid 174 to amino acid 183 of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone bp783_3 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 44.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 99 to nucleotide 1514;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 171 to nucleotide 1514;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 57 to nucleotide 623;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bu45_2 deposited under accession number ATCC 98369;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bu45_2 deposited under accession number ATCC 98369;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bu45_2 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bu45_2 deposited under accession number ATCC 98369;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 231 to amino acid 240 of SEQ ID NO:4;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 99 to nucleotide 1514; the nucleotide sequence of SEQ ID NO:3 from nucleotide 171 to nucleotide 1514; the nucleotide sequence of SEQ ID NO:3 from nucleotide 57 to nucleotide 623; the nucleotide sequence of the full-length protein coding sequence of clone bu45_2 deposited under accession number ATCC 98369; or the nucleotide sequence of a mature protein coding sequence of clone bu45_2 deposited under accession number ATCC 98369. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone bu45_2 deposited under accession number ATCC 98369. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 175.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 175;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising the amino acid sequence from amino acid 231 to amino acid 240 of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone bu45_2 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 175.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 87 to nucleotide 980;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 147 to nucleotide 980;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ct864_4 deposited under accession number ATCC 98369;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ct864_4 deposited under accession number ATCC 98369;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ct864_4 deposited under accession number ATCC 98369;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ct864_4 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 144 to amino acid 153 of SEQ ID NO:6;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 87 to nucleotide 980; the nucleotide sequence of SEQ ID NO:5 from nucleotide 147 to nucleotide 980; the nucleotide sequence of the full-length protein coding sequence of clone ct864_4 deposited under accession number ATCC 98369; or the nucleotide sequence of a mature protein coding sequence of clone ct864_4 deposited under accession number ATCC 98369. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone ct864_4 deposited under accession number ATCC 98369. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6 from amino acid 189 to amino acid 290.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:5.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 189 to amino acid 290;

(c) fragments of the amino acid sequence of SEQ ID NO:6 comprising the amino acid sequence from amino acid 144 to amino acid 153 of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone ct864_4 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6 or the amino acid sequence of SEQ ID NO:6 from amino acid 189 to amino acid 290.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 242 to nucleotide 580;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 1 to nucleotide 387;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone df396_1 deposited under accession number ATCC 98369;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone df396_1 deposited under accession number ATCC 98369;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone df396_1 deposited under accession number ATCC 98369;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone df396_1 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 51 to amino acid 60 of SEQ ID NO:8;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:7 from nucleotide 242 to nucleotide 580; the nucleotide sequence of SEQ ID NO:7 from nucleotide 1 to nucleotide 387; the nucleotide sequence of the full-length protein coding sequence of clone df396_1 deposited under accession number ATCC 98369; or the nucleotide sequence of a mature protein coding sequence of clone df396_1 deposited under accession number ATCC 98369. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone df396_1 deposited under accession number ATCC 98369. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 48.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:7.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 48;

(c) fragments of the amino acid sequence of SEQ ID NO:8 comprising the amino acid sequence from amino acid 51 to amino acid 60 of SEQ ID NO:8; and

(d) the amino acid sequence encoded by the cDNA insert of clone df396_1 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:8 or the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 48.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 236 to nucleotide 1213;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 1386 to nucleotide 1833;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dh1135_9 deposited under accession number ATCC 98369;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dh1135_9 deposited under accession number ATCC 98369;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dh1135_9 deposited under accession number ATCC 98369;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dh1135_9 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 157 to amino acid 166 of SEQ ID NO:10;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:9 from nucleotide 236 to nucleotide 1213; the nucleotide sequence of SEQ ID NO:9 from nucleotide 1386 to nucleotide 1833; the nucleotide sequence of the full-length protein coding sequence of clone dh1135_9 deposited under accession number ATCC 98369; or the nucleotide sequence of a mature protein coding sequence of clone dh1135_9 deposited under accession number ATCC 98369. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone dh1135_9 deposited under accession number ATCC 98369. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein

comprising the amino acid sequence of SEQ ID NO:31 from amino acid 1 to amino acid 147.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:9.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) the amino acid sequence of SEQ ID NO:31 from amino acid 1 to amino acid 147;

(c) fragments of the amino acid sequence of SEQ ID NO:10 comprising the amino acid sequence from amino acid 157 to amino acid 166 of SEQ ID NO:10; and

(d) the amino acid sequence encoded by the cDNA insert of clone dh1135_9 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:10 or the amino acid sequence of SEQ ID NO:31 from amino acid 1 to amino acid 147.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 334 to nucleotide 675;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 409 to nucleotide 675;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dn809_5 deposited under accession number ATCC 98369;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dn809_5 deposited under accession number ATCC 98369;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dn809_5 deposited under accession number ATCC 98369;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dn809_5 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising the amino acid sequence from amino acid 52 to amino acid 61 of SEQ ID NO:12;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:11 from nucleotide 334 to nucleotide 675; the nucleotide sequence of SEQ ID NO:11 from nucleotide 409 to nucleotide 675; the nucleotide sequence of the full-length protein coding sequence of clone dn809_5 deposited under accession number ATCC 98369; or the nucleotide sequence of a mature protein coding sequence of clone dn809_5 deposited under accession number ATCC 98369. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone dn809_5 deposited under accession number ATCC 98369. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 110.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:11.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:12;

(b) the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 110;

(c) fragments of the amino acid sequence of SEQ ID NO:12 comprising the amino acid sequence from amino acid 52 to amino acid 61 of SEQ ID NO:12; and

(d) the amino acid sequence encoded by the cDNA insert of clone dn809_5 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:12 or the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 110.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 447 to nucleotide 791;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 597 to nucleotide 791;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1 to nucleotide 546;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ej224_1 deposited under accession number ATCC 98369;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ej224_1 deposited under accession number ATCC 98369;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ej224_1 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ej224_1 deposited under accession number ATCC 98369;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising the amino acid sequence from amino acid 52 to amino acid 61 of SEQ ID NO:14;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:13 from nucleotide 447 to nucleotide 791; the nucleotide sequence of SEQ ID NO:13 from nucleotide 597 to nucleotide 791; the nucleotide sequence of SEQ ID NO:13 from nucleotide 1 to nucleotide 546; the nucleotide sequence of the full-length protein coding sequence of clone ej224_1 deposited under accession number ATCC 98369; or the nucleotide sequence of a mature protein coding sequence of clone ej224_1 deposited under accession number ATCC 98369. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone ej224_1 deposited under accession number ATCC 98369. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14 from amino acid 82 to amino acid 100.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:13.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:14;

(b) the amino acid sequence of SEQ ID NO:14 from amino acid 82 to amino acid 100;

(c) fragments of the amino acid sequence of SEQ ID NO:14 comprising the amino acid sequence from amino acid 52 to amino acid 61 of SEQ ID NO:14; and

(d) the amino acid sequence encoded by the cDNA insert of clone ej224_1 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:14 or the amino acid sequence of SEQ ID NO:14 from amino acid 82 to amino acid 100.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 18 to nucleotide 347;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 1 to nucleotide 345;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ek591_1 deposited under accession number ATCC 98369;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ek591_1 deposited under accession number ATCC 98369;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ek591_1 deposited under accession number ATCC 98369;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ek591_1 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising the amino acid sequence from amino acid 50 to amino acid 59 of SEQ ID NO:16;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:15 from nucleotide 18 to nucleotide 347; the nucleotide sequence of SEQ ID NO:15 from nucleotide 1 to nucleotide 345; the nucleotide sequence of the full-length protein coding sequence of clone ek591_1 deposited under accession number ATCC 98369; or the

nucleotide sequence of a mature protein coding sequence of clone ek591_1 deposited under accession number ATCC 98369. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone ek591_1 deposited under accession number ATCC 98369. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 109.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:15.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 109;

(c) fragments of the amino acid sequence of SEQ ID NO:16 comprising the amino acid sequence from amino acid 50 to amino acid 59 of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone ek591_1 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:16 or the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 109.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 593 to nucleotide 1663;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 833 to nucleotide 1663;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 648 to nucleotide 1063;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone er381_1 deposited under accession number ATCC 98369;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone er381_1 deposited under accession number ATCC 98369;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone er381_1 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone er381_1 deposited under accession number ATCC 98369;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising the amino acid sequence from amino acid 173 to amino acid 182 of SEQ ID NO:18;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:17 from nucleotide 593 to nucleotide 1663; the nucleotide sequence of SEQ ID NO:17 from nucleotide 833 to nucleotide 1663; the nucleotide sequence of SEQ ID NO:17 from nucleotide 648 to nucleotide 1063; the nucleotide sequence of the full-length protein coding sequence of clone er381_1 deposited under accession number ATCC 98369; or the nucleotide sequence of a mature protein coding sequence of clone er381_1 deposited under accession number ATCC 98369. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone er381_1 deposited under accession number ATCC 98369. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18 from amino acid 20 to amino acid 157.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:17.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:18;

(b) the amino acid sequence of SEQ ID NO:18 from amino acid 20 to amino acid 157;

(c) fragments of the amino acid sequence of SEQ ID NO:18 comprising the amino acid sequence from amino acid 173 to amino acid 182 of SEQ ID NO:18; and

(d) the amino acid sequence encoded by the cDNA insert of clone er381_1 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:18 or the amino acid sequence of SEQ ID NO:18 from amino acid 20 to amino acid 157.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 1055 to nucleotide 1246;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 759 to nucleotide 1152;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone gq38_1 deposited under accession number ATCC 98369;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone gq38_1 deposited under accession number ATCC 98369;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone gq38_1 deposited under accession number ATCC 98369;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone gq38_1 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment comprising the amino acid sequence from amino acid 20 to amino acid 29 of SEQ ID NO:20;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:19 from nucleotide 1055 to nucleotide 1246; the nucleotide sequence of SEQ ID NO:19 from nucleotide 759 to nucleotide 1152; the nucleotide sequence of the full-length protein coding sequence of clone gq38_1 deposited under accession number ATCC 98369; or the nucleotide sequence of a mature protein coding sequence of clone gq38_1 deposited under accession number ATCC 98369. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone gq38_1 deposited under accession number ATCC 98369. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 32.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:19.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:20;

(b) the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 32;

(c) fragments of the amino acid sequence of SEQ ID NO:20 comprising the amino acid sequence from amino acid 20 to amino acid 29 of SEQ ID NO:20; and

(d) the amino acid sequence encoded by the cDNA insert of clone gq38_1 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:20 or the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 32.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions. Also provided by the present invention are organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

(a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and

(b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

DETAILED DESCRIPTION

ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide

sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature forms) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g. , receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

Clone "bp783_3"

A polynucleotide of the present invention has been identified as clone "bp783_3". bp783_3 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bp783_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bp783_3 protein").

The nucleotide sequence of bp783_3 as presently determined is reported in SEQ ID NO:1. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bp783_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bp783_3 should be approximately 2300 bp.

The nucleotide sequence disclosed herein for bp783_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bp783_3 demonstrated at least some similarity with sequences identified as AA099506 (zm17b06.r1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 525875 5'), AA703257 (zi70f10.s1 Soares fetal liver spleen 1NFLS S1 Homo sapiens

cDNA clone 436171 3'), N33318 (yy08a03.s1 Homo sapiens cDNA clone 270604 3'), N35074 (yy19b06.s1 Homo sapiens cDNA clone 271667 3'), and W29359 (mb96f10.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 337291 5'). Based upon sequence similarity, bp783_3 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of bp783_3 indicates that it may contain a GAAA repeat sequence.

Clone "bu45_2"

A polynucleotide of the present invention has been identified as clone "bu45_2". bu45_2 was isolated from a human adult placenta cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bu45_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bu45_2 protein").

The nucleotide sequence of bu45_2 as presently determined is reported in SEQ ID NO:3. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bu45_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4. Amino acids 12 to 24 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 25, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bu45_2 should be approximately 1850 bp.

The nucleotide sequence disclosed herein for bu45_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bu45_2 demonstrated at least some similarity with sequences identified as AA041196 (zf09e05.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 376448 3'), AA452391 (zx29c10.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone 787890 5'), Q61260 (Human brain Expressed Sequence Tag EST01280), R13864 (yf65e05.r1 Homo sapiens cDNA clone 27004 5'), and R18560 (yf95b10.r1 Homo sapiens cDNA clone 30142 5). The predicted amino acid sequence disclosed herein for bu45_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted bu45_2 protein demonstrated at least some similarity to sequences identified as R99416 (Aminopeptidase precursor of Aeromonas

caviae). Based upon sequence similarity, bu45_2 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts three additional potential transmembrane domains within the bu45_2 protein sequence, centered around amino acids 137, 205, and 456 of SEQ ID NO:4, respectively.

Clone "ct864_4"

A polynucleotide of the present invention has been identified as clone "ct864_4". ct864_4 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. ct864_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ct864_4 protein").

The nucleotide sequence of ct864_4 as presently determined is reported in SEQ ID NO:5. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the ct864_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6. Amino acids 8 to 20 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 21, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ct864_4 should be approximately 1150 bp.

The nucleotide sequence disclosed herein for ct864_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. ct864_4 demonstrated at least some similarity with sequences identified as AA725566 (ai24d02.s1 Soares testis NHT Homo sapiens cDNA clone 1343715 3' similar to TR Q99795 Q99795 A33 ANTIGEN PRECURSOR), N90730 (za90e09.s1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 299848 3'), T89217 (ye12c02.r1 Homo sapiens cDNA clone 117506 5'), and W80145 (me91g01.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone 402960 5'). The predicted amino acid sequence disclosed herein for ct864_4 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted ct864_4 protein demonstrated at least some similarity to sequences identified as U79725 (A33 antigen precursor [Homo sapiens]). A33 antigen precursor is a transmembrane protein and a member of the immunoglobulin superfamily (Heath *et al.*, 1997, *Proc. Natl. Acad. Sci. USA*

94: 469-474). Based upon sequence similarity, ct864_4 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domains within the ct864_4 protein sequence centered around amino acid 247 of SEQ ID NO:6.

Clone "df396_1"

A polynucleotide of the present invention has been identified as clone "df396_1". df396_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. df396_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "df396_1 protein").

The nucleotide sequence of df396_1 as presently determined is reported in SEQ ID NO:7. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the df396_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:8.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone df396_1 should be approximately 2500 bp.

The nucleotide sequence disclosed herein for df396_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. df396_1 demonstrated at least some similarity with sequences identified as T69764 (yd14c05.s1 Homo sapiens cDNA clone 108200 3') and Z80897 (Human DNA sequence from cosmid E132D12 on chromosome 22q12-qter). Based upon sequence similarity, df396_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts two potential transmembrane domains within the df396_1 protein sequence, centered around amino acids 40 and 80 of SEQ ID NO:8, respectively.

Clone "dh1135_9"

A polynucleotide of the present invention has been identified as clone "dh1135_9". dh1135_9 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer

analysis of the amino acid sequence of the encoded protein. dh1135_9 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "dh1135_9 protein").

The nucleotide sequence of dh1135_9 as presently determined is reported in SEQ ID NO:9. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the dh1135_9 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:10. Another potential dh1135_9 reading frame and predicted amino acid sequence is encoded by basepairs 1394 to 1879 of SEQ ID NO:9 and is reported in SEQ ID NO:31. Amino acids 84 to 96 of SEQ ID NO:31 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 97, or are a transmembrane domain. The open reading frames of SEQ ID NO:10 and SEQ ID NO:31 could be joined if one or more frameshifts were introduced into the nucleotide sequence of SEQ ID NO:9 between basepairs 1000 and 1400.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone dh1135_9 should be approximately 2000 bp.

The nucleotide sequence disclosed herein for dh1135_9 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. dh1135_9 demonstrated at least some similarity with sequences identified as AA102652 (zn73b01.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone 563785 3'), AA207179 (zq73b05.r1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA clone 647217 5'), AA233641 (zr43f02.r1 Soares NhHMPu S1 Homo sapiens cDNA clone 666171 5' similar to TR:G1109804 G1109804 CODED FOR BY C. ELEGANS CDNA CEESW58F), AA238618 (my33e04.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA clone 697662 5'), AA588137 (nm99a06.s1 NCI_CGAP_Co9 Homo sapiens cDNA clone IMAGE:1076338), W40329 (zc81c12.r1 Pancreatic Islet Homo sapiens cDNA clone 328726 5'), and W45396 (zc81c12.s1 Pancreatic Islet Homo sapiens cDNA clone 328726 3'). The predicted amino acid sequence disclosed herein for dh1135_9 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted dh1135_9 protein demonstrated at least some similarity to sequences identified as U41531 (coded for by C. elegans cDNA CEESW58F [Caenorhabditis elegans]). Based upon sequence similarity, dh1135_9 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts two potential transmembrane domains within

the dh1135_9 protein sequence of SEQ ID NO:10, one around amino acid 50 and another around amino acid 280 of SEQ ID NO:10.

Clone "dn809_5"

A polynucleotide of the present invention has been identified as clone "dn809_5". dn809_5 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. dn809_5 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "dn809_5 protein").

The nucleotide sequence of dn809_5 as presently determined is reported in SEQ ID NO:11. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the dn809_5 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:12. Amino acids 13 to 25 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 26, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone dn809_5 should be approximately 1000 bp.

The nucleotide sequence disclosed herein for dn809_5 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. dn809_5 demonstrated at least some similarity with sequences identified as AA252421 (zs13a07.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone 685044 5'), AA400027 (zu68f11.r1 Soares testis NHT Homo sapiens cDNA clone 743181 5' similar to contains element MSR1 repetitive element), T79197 (yd70f07.s1 Homo sapiens cDNA clone 113605 3'), and T79284 (yd70f07.r1 Homo sapiens cDNA clone 113605 5'). Based upon sequence similarity, dn809_5 proteins and each similar protein or peptide may share at least some activity.

Clone "ej224_1"

A polynucleotide of the present invention has been identified as clone "ej224_1". ej224_1 was isolated from a human adult placenta cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer

analysis of the amino acid sequence of the encoded protein. ej224_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ej224_1 protein").

The nucleotide sequence of ej224_1 as presently determined is reported in SEQ ID NO:13. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the ej224_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:14. Amino acids 38 to 50 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 51, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ej224_1 should be approximately 2300 bp.

The nucleotide sequence disclosed herein for ej224_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. ej224_1 demonstrated at least some similarity with sequences identified as H79156 (yu47a04.r1 Homo sapiens cDNA clone 229230 5' similar to contains Alu repetitive element), M87922 (Human carcinoma cell-derived Alu RNA transcript, clone CD139), and N64587 (yz51h09.s1 Homo sapiens cDNA clone 286625 3' similar to contains Alu repetitive element). Based upon sequence similarity, ej224_1 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of ej224_1 indicates that it may contain an Alu repetitive element.

Clone "ek591_1"

A polynucleotide of the present invention has been identified as clone "ek591_1". ek591_1 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. ek591_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ek591_1 protein").

The nucleotide sequence of ek591_1 as presently determined is reported in SEQ ID NO:15. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the ek591_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:16. Another potential ek591_1 reading frame and predicted amino acid sequence is encoded by basepairs 351 to 599 of SEQ ID

NO:15 and is reported in SEQ ID NO:32; the TopPredII computer program predicts a potential transmembrane domain within the SEQ ID NO:32 amino acid sequence. If the stop codon at basepairs 348-350 of SEQ ID NO:15 were altered to encode an amino acid, the open reading frame of SEQ ID NO:16 would be joined to that of SEQ ID NO:32.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ek591_1 should be approximately 1300 bp.

The nucleotide sequence disclosed herein for ek591_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. ek591_1 demonstrated at least some similarity with sequences identified as AA149073 (zl45d10.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 504883 5' similar to TR G1230697 G1230697 CHROMOSOME XVI COSMID 9513), AA149074 (zl45d10.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 504883 3'), U51033 (Saccharomyces cerevisiae chromosome XVI cosmid 9513), and W31137 (zb45g03.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 306580 5'). The predicted amino acid sequence disclosed herein for ek591_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted ek591_1 protein demonstrated at least some similarity to sequences identified as U51033 (P9513.2 gene product [Saccharomyces cerevisiae]). Based upon sequence similarity, ek591_1 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of ek591_1 indicates that it may contain repetitive elements.

Clone "er381_1"

A polynucleotide of the present invention has been identified as clone "er381_1". er381_1 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. er381_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "er381_1 protein").

The nucleotide sequence of er381_1 as presently determined is reported in SEQ ID NO:17. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the er381_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:18. Amino acids 68 to 80 are a predicted

leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 81, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone er381_1 should be approximately 2200 bp.

The nucleotide sequence disclosed herein for er381_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. er381_1 demonstrated at least some similarity with sequences identified as AA043260 (zk49g05.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 486200 3'), AA385070 (EST98667 Thyroid Homo sapiens cDNA 5' end), H28240 (yl60b04.r1 Homo sapiens cDNA clone 162607 5'), H28273 (yl60h04.r1 Homo sapiens cDNA clone 162679 5'), T23745 (Human gene signature HUMGS05632), W29691 (mc07h04.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 347863 5'), and W97088 (mf61d08.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone 418767 5'). Based upon sequence similarity, er381_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts two potential transmembrane domains within the er381_1 protein sequence, one around amino acid 200 and another around amino acid 220 of SEQ ID NO:18. The nucleotide sequence of er381_1 indicates that it may contain a TAR1 repetitive element.

Clone "gq38_1"

A polynucleotide of the present invention has been identified as clone "gq38_1". gq38_1 was isolated from a human adult pineal gland cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. gq38_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "gq38_1 protein").

The nucleotide sequence of gq38_1 as presently determined is reported in SEQ ID NO:19. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the gq38_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:20.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone gq38_1 should be approximately 1500 bp.

The nucleotide sequence disclosed herein for gq38_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. gq38_1 demonstrated at least some similarity with sequences identified as AA134939 (zo26b06.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 587987 3'), AA195485 (zp87h08.s1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone 627231 3'), AA280722 (zs96e09.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone 711496 5'), H85699 (ys68e04.r1 Homo sapiens cDNA clone 219966 5' similar to contains Alu repetitive element), N98571 (za69g01.r1 Homo sapiens cDNA clone 297840 5'), R81264 (yj01a02.r1 Homo sapiens cDNA clone 147434 5'), and W76442 (zd61b07.r1 Soares fetal heart). Based upon sequence similarity, gq38_1 proteins and each similar protein or peptide may share at least some activity.

Deposit of Clones

Clones bp783_3, bu45_2, ct864_4, df396_1, dh1135_9, dn809_5, ej224_1, ek591_1, er381_1, and gq38_1 were deposited on March 21, 1997 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98369, from which each clone comprising a particular polynucleotide is obtainable. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b).

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Fig. 1. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* **19**: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* **9**: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for

expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of the oligonucleotide probe that was used to isolate each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

Clone	Probe Sequence
bp783_3	SEQ ID NO:21
bu45_2	SEQ ID NO:22
ct864_4	SEQ ID NO:23
df396_1	SEQ ID NO:24
dh1135_9	SEQ ID NO:25
dn809_5	SEQ ID NO:26
ej224_1	SEQ ID NO:27
ek591_1	SEQ ID NO:28
er381_1	SEQ ID NO:29
gq38_1	SEQ ID NO:30

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoaramidite residue rather than a nucleotide (such as , for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramadite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

(a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;

(b) It should be designed to have a T_m of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

The oligonucleotide should preferably be labeled with g-^{32}P ATP (specific activity 6000 Ci/mmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately 4e+6 dpm/pmole.

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100 µl of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100 µg/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100 µg/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100 µg/ml of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to 1e+6 dpm/mL. The filter is then preferably incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis, hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, Bio/Technology 10, 773-778 (1992) and in R.S. McDowell, *et al.*, J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form(s) of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence(s) of the mature form(s) of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The

desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* **15(7)**: 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* **62(1)**: 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* **58**: 1-39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* **14(9)**: 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* **90(16)**: 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* **91(2)**: 719-722; all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* **336**: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614,396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologues of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide. Preferably, polynucleotide species homologues have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, and protein species homologues have at least 30% sequence identity (more preferably, at least 45% identity; most preferably at least 60% identity) with the given protein, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides or the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Species homologues may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species. Preferably, species homologues are those isolated from mammalian species. Most preferably, species homologues are those isolated from certain mammalian species such as, for example, *Pan troglodytes, Gorilla gorilla, Pongo pygmaeus, Hylobates concolor, Macaca mulatta, Papio papio, Papio hamadryas, Cercopithecus aethiops, Cebus capucinus, Aotus trivirgatus, Sanguinus oedipus, Microcebus murinus, Mus musculus, Rattus norvegicus, Cricetulus griseus, Felis catus, Mustela vison, Canis familiaris, Oryctolagus cuniculus, Bos taurus, Ovis aries, Sus scrofa,* and *Equus caballus,* for which genetic maps have been created allowing the identification of syntenic relationships between the genomic organization of genes in one species and the genomic organization of the related genes in another species

(O'Brien and Seuánez, 1988, *Ann. Rev. Genet.* **22**: 323-351; O'Brien *et al.*, 1993, *Nature Genetics* 3:103-112; Johansson *et al.*, 1995, *Genomics* **25**: 682-690; Lyons *et al.*, 1997, *Nature Genetics* **15**: 47-56; O'Brien *et al.*, 1997, *Trends in Genetics* **13(10)**: 393-399; Carver and Stubbs, 1997, *Genome Research* 7:1123-1137; all of which are incorporated by reference herein).

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotides which also encode proteins which are identical or have significantly similar sequences to those encoded by the disclosed polynucleotides. Preferably, allelic variants have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps. Allelic variants may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from individuals of the appropriate species.

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp)‡	Hybridization Temperature and Buffer†	Wash Temperature and Buffer†
A	DNA:DNA	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	<50	T_B*; 1xSSC	T_B*; 1xSSC
C	DNA:RNA	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	<50	T_D*; 1xSSC	T_D*; 1xSSC
E	RNA:RNA	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	<50	T_F*; 1xSSC	T_F*; 1xSSC
G	DNA:DNA	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	<50	T_H*; 4xSSC	T_H*; 4xSSC
I	DNA:RNA	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	<50	T_J*; 4xSSC	T_J*; 4xSSC
K	RNA:RNA	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	<50	T_L*; 2xSSC	T_L*; 2xSSC
M	DNA:DNA	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	<50	T_N*; 6xSSC	T_N*; 6xSSC
O	DNA:RNA	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	<50	T_P*; 6xSSC	T_P*; 6xSSC
Q	RNA:RNA	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	<50	T_R*; 4xSSC	T_R*; 4xSSC

‡: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

†: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

*T_B - T_R: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6($\log_{10}$[Na^+]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na^+] is the concentration of sodium ions in the hybridization buffer ([Na^+] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium*, or any bacterial

strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant

methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which

the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays

for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., J. Immunol. 149:3778-3783, 1992; Bowman et al., J. Immunol. 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ, Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for

example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or

tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, *89*:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of

viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface.

Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (*e.g.*, B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowmanet al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro* antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter

7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent

myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc.., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland,

H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of

congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and

Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the

first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used

to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or caricadic cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic

lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunolgobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be

administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred

pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1ng to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal

antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-

aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorbtion of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in

the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Racie, Lisa A.
Merberg, David
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.

(ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(iii) NUMBER OF SEQUENCES: 32

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Genetics Institute, Inc.
(B) STREET: 87 CambridgePark Drive
(C) CITY: Cambridge
(D) STATE: MA
(E) COUNTRY: U.S.A.
(F) ZIP: 02140

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Sprunger, Suzanne A.
(B) REGISTRATION NUMBER: 41,323

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 498-8284
(B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2199 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGTTGGCAGG TGGAGAGGCA GGTTGGGAGG GAAAGTCGGG GGAGGACGCG GAAGAGGAGC 60

TGTGGGAAGG GGGAGGAGGG AGGGAGGAAA AGAGGAGGAG GCGGAGGAGA ACTGAGCAGA 120

GCAGAGCATC GAGCCAAAGG GGAGATGAGT TTGTCTGTCC TCTGCTGAGG CTACGGCCGG 180

GCCTAGGGAA CTGGGAGCTT GGGTGGAAGC GACACCCGTG GAAGTGGGAG GAGGTGGCGC 240

CGGGACTTTA ACCCCTTGTG GGCTCTGCGG CAGGGGATTT AACCCTTTGT GGATCTGGCC 300

CCTCGGAGGC AGCGTCATCG GTAGTTTTAA CCCCTTCGGG GCTGGGTTTC ACGCACTGGA 360

CTTACCCTCA TCACCTTGCT CACCAACTCC TTTATTGGGG TGCTCCGCTT GGAGGTTTGA 420

GGCCCACCTC CGCCCATTAC GTACTGTTCC TGCCGCTGCA CCCCCTTGGA CCCGCTAGCT 480

GGCCGCACTG TGGGCGCTTA ACCCTTTACT GACTTGAGCT CCCCAGATTG CAGTTGGAGT 540

TTGCTGATAG AAGGACTAGC TAAAGGCGTC ACTGCAGGAA TTACAAACTG AAGAGGACTC 600

TGTTGGACTG TTTTTTTTTT CTTTTTCTTT TTTTTAAGAA AAACCCATTT TTTTCCTTAA 660

GGACTTACTA GCCAAAATTT CTTAAACTTC GAGGACTCTA CTAGCCATGG CCGAGCCATT 720

CTTGTCAGAA TATCAACACC AGCCTCAAAC TAGCAACTGT ACAGGTGCTG CTGCTGTCCA 780

GGAAGAGCTG AACCCTGAGC GCCCCCCAGG CGCGGAGGAG CGGGTGCCCG AGGAGGACAG 840

TAGGTGGCAA TCGAGAGCGT TCCCCCAGTT GGGTGGCCGT CCGGGGCCGG AGGGGGAAGG 900

GAGCCTGGAA TCCCAACCAC CTCCCTTGCA GACCCAGGCC TGTCCAGAAT CTAGCTGCCT 960

GAGAGAGGGC GAGAAGGGCC AGAATGGGGA CGACTCGTCC GCTGGCGGCG ACTTCCCGCC 1020

GCCGGCAGAA GTGGAACCGA CGCCCGAGGC CGAGCTGCTC GCCCAGCCTT GTCATGACTC 1080

CGAGGCCAGT AAGTTGGGGG CTCCTGCCGC AGGGGGCGAA GAGGAGTGGG GACAGCAGCA 1140

GAGACAGCTG GGGAAGAAAA AACATAGGAG ACGCCCGTCC AAGAAGAAGC GGCATTGGAA 1200

ACCGTACTAC AAGCTGACCT GGGAAGAGAA GAAAAAGTTC GACGAGAAAC AGAGCCTTCG 1260

AGCTTCAAGG ATCCGAGCCG AGATGTTCGC CAAGGGCCAG CCGGTCGCGC CCTATAACAC 1320

CACGCAGTTC CTCATGGATG ATCACGACCA GGAGGAGCCG GATCTCAAAA CCGGCCTGTA 1380

CTCCAAGCGG GCCGCCGCCA AATCCGACGA CACCAGCGAT GACGACTTCA TGGAAGAAGG 1440

GGGTGAGGAG GATGGGGGCA GCGATGGGAT GGGAGGGGAC GGCAGCGAGT TTCTGCAGCG 1500

GGACTTCTCG GAGACGTACG AGCGGTACCA CACGGAGAGC CTGCAGAACA TGAGCAAGCA 1560

GGAGCTCATC AAGGAGTACC TGGAACTGGA GAAGTGCCTC TCGCGCATGG AGGACGAGAA 1620

CAACCGGCTG CGGCTGGAGA GCAAGCGGCT GGGTGGCGAC GACGCGCGTG TGCGGGAGCT 1680

GGAGCTGGAG CTGGACCGGC TGCGCGCCGA GAACCTCCAG CTGCTGACCG AGAACGAACT 1740

GCACCGGCAG CAGGAGCGAG CGCCGCTTTC CAAGTTTGGA GACTAGACTG AAACTTTTTT 1800

GGGGGAGGGG GCAAAGGGGA CTTTTTACAG TGATGGAATG TAACATTATA TACATGTGTA 1860

TATAAGACAG TGGACCTTTT TATGACACAT AATCAGAAGA GAAATCCCCC TGGCTTTGGT 1920

TGGTTTCGTA AATTTAGCTA TATGTAGCTT GCGTGCTTTC TCCTGTTCTT TTAATTATGT 1980

GAAACTGAAG AGTTGCTTTT CTTGTTTTCC TTTTTAGAAG TTTTTTTCCT TAATGTGAAA 2040

GTAATTTGAC CAAGTTATAA TGCATTTTTG TTTTTAACAA ATCCCCTCCT TAAACGGAGC 2100

TATAAGGTGG CCAAATCTGA GAACAATTAA ATTCATTTTA GTTATAATAA ATTTAATATT 2160

TGTAAATGTA AAAAAAAAAA AAAAAAAAAA AAAAAAAAA 2199

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 359 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Glu Pro Phe Leu Ser Glu Tyr Gln His Gln Pro Gln Thr Ser
1 5 10 15

Asn Cys Thr Gly Ala Ala Ala Val Gln Glu Glu Leu Asn Pro Glu Arg
20 25 30

Pro Pro Gly Ala Glu Glu Arg Val Pro Glu Glu Asp Ser Arg Trp Gln
35 40 45

Ser Arg Ala Phe Pro Gln Leu Gly Gly Arg Pro Gly Pro Glu Gly Glu
50 55 60

Gly Ser Leu Glu Ser Gln Pro Pro Pro Leu Gln Thr Gln Ala Cys Pro
65 70 75 80

Glu Ser Ser Cys Leu Arg Glu Gly Glu Lys Gly Gln Asn Gly Asp Asp
85 90 95

Ser Ser Ala Gly Gly Asp Phe Pro Pro Pro Ala Glu Val Glu Pro Thr
100 105 110

Pro Glu Ala Glu Leu Leu Ala Gln Pro Cys His Asp Ser Glu Ala Ser
115 120 125

Lys Leu Gly Ala Pro Ala Ala Gly Gly Glu Glu Glu Trp Gly Gln Gln
130 135 140

Gln Arg Gln Leu Gly Lys Lys Lys His Arg Arg Arg Pro Ser Lys Lys
145 150 155 160

Lys Arg His Trp Lys Pro Tyr Tyr Lys Leu Thr Trp Glu Glu Lys Lys
165 170 175

Lys Phe Asp Glu Lys Gln Ser Leu Arg Ala Ser Arg Ile Arg Ala Glu
180 185 190

Met Phe Ala Lys Gly Gln Pro Val Ala Pro Tyr Asn Thr Thr Gln Phe
195 200 205

Leu Met Asp Asp His Asp Gln Glu Glu Pro Asp Leu Lys Thr Gly Leu
210 215 220

Tyr Ser Lys Arg Ala Ala Ala Lys Ser Asp Asp Thr Ser Asp Asp Asp
225 230 235 240

Phe Met Glu Glu Gly Gly Glu Glu Asp Gly Gly Ser Asp Gly Met Gly
245 250 255

Gly Asp Gly Ser Glu Phe Leu Gln Arg Asp Phe Ser Glu Thr Tyr Glu
260 265 270

Arg Tyr His Thr Glu Ser Leu Gln Asn Met Ser Lys Gln Glu Leu Ile
275 280 285

Lys Glu Tyr Leu Glu Leu Glu Lys Cys Leu Ser Arg Met Glu Asp Glu
290 295 300

Asn Asn Arg Leu Arg Leu Glu Ser Lys Arg Leu Gly Gly Asp Asp Ala
305 310 315 320

Arg Val Arg Glu Leu Glu Leu Glu Leu Asp Arg Leu Arg Ala Glu Asn
325 330 335

Leu Gln Leu Leu Thr Glu Asn Glu Leu His Arg Gln Gln Glu Arg Ala
340 345 350

Pro Leu Ser Lys Phe Gly Asp
355

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1851 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGCTAGGCCG CGAGCTTAGT CCTGGGAGCC GCCTCCGTCG CCGCCGTCAG AGCCGCCCTA 60

TCAGATTATC TTAACAAGAA AACCAACTGG AAAAAAAAAT GAAATTCCTT ATCTTCGCAT 120

TTTTCGGTGG TGTTCACCTT TTATCCCTGT GCTCTGGGAA AGCTATATGC AAGAATGGCA 180

TCTCTAAGAG GACTTTTGAA GAAATAAAAG AAGAAATAGC CAGCTGTGGA GATGTTGCTA 240

AAGCAATCAT CAACCTAGCT GTTTATGGTA AAGCCCAGAA CAGATCCTAT GAGCGATTGG 300

CACTTCTGGT TGATACTGTT GGACCCAGAC TGAGTGGCTC CAAGAACCTA GAAAAAGCCA 360

TCCAAATTAT GTACCAAAAC CTGCAGCAAG ATGGGCTGGA GAAAGTTCAC CTGGAGCCAG 420

TGAGAATACC CCACTGGGAG AGGGGAGAAG AATCAGCTGT GATGCTGGAG CCAAGAATTC 480

ATAAGATAGC CATCCTGGGT CTTGGCAGCA GCATTGGGAC TCCTCCAGAA GGCATTACAG 540

CAGAAGTTCT GGTGGTGACC TCTTTCGATG AACTGCAGAG AAGGGCCTCA GAAGCAAGAG 600

GGAAGATTGT TGTTTATAAC CAACCTTACA TCAACTACTC AAGGACGGTG CAATACCGAA 660

CGCAGGGGGC GGTGGAAGCT GCCAAGGTGG GGGCTTTGGC ATCTCTCATT CGATCCGTGG 720

CCTCCTTCTC CATCTACAGT CCTCACACAG GTATTCAGGA ATACCAGGAT GGCGTGCCCA 780

AAATTCCAAC AGCCTGTATT ACGGTGGAAG ATGCAGAAAT GATGTCAAGA ATGGCTTCTC 840

ATGGGATCAA AATTGTCATT CAGCTAAAGA TGGGGGCAAA GACCTACCCA GATACTGATT 900

CCTTCAACAC TGTAGCAGAG ATCACTGGGA GCAAATATCC AGAACAGGTT GTACTGGTCA 960

GTGGACATCT GGACAGCTGG GATGTTGGGC AGGGTGCCAT GGATGATGGC GGTGGAGCCT 1020

TTATATCATG GGAAGCACTC TCACTTATTA AAGATCTTGG GCTGCGTCCA AAGAGGACTC 1080

TGCGGCTGGT GCTCTGGACT GCAGAAGAAC AAGGTGGAGT TGGTGCCTTC CAGTATTATC 1140

AGTTACACAA GGTAAATATT TCCAACTACA GTCTGGTGAT GGAGTCTGAC GCAGGAACCT 1200

TCTTACCCAC TGGGCTGCAA TTCACTGGCA GTGAAAAGGC CAGGGCCATC ATGGAGGAGG 1260

TTATGAGCCT GCTGCAGCCC CTCAATATCA CTCAGGTCCT GAGCCATGGA GAAGGGACAG 1320

ACATCAACTT TTGGATCCAA GCTGGAGTGC CTGGAGCCAG TCTACTTGAT GACTTATACA 1380

AGTATTTCTT CTTCCATCAC TCCCACGGAG ACACCATGAC TGTCATGGAT CCAAAGCAGA 1440

TGAATGTTGC TGCTGCTGTT TGGGCTGTTG TTTCTTATGT TGTTGCAGAC ATGGAAGAAA 1500

TGCTGCCTAG GTCCTAGAAA CAGTAAGAAA GAAACGTTTT CATGCTTCTG GCCAGGAATC 1560

CTGGGTCTGC AACTTTGGAA AACTCCTCTT CACATAACAA TTTCATCCAA TTCATCTTCA 1620

AAGCACAACT CTATTTCATG CTTTCTGTTA TTATCTTTCT TGATACTTTC CAAATTCTCT 1680

GATTCTAGAA AAAGGAATCA TTCTCCCCTC CCTCCCACCA CATAGAATCA ACATATGGTA 1740

GGGATTACAG TGGGGGCATT TCTTTATATC ACCTCTTAAA AACATTGTTT CCACTTTAAA 1800

AGTAAACACT TAATAAATTT TTGGAAGATC TCTGAAAAAA AAAAAAAAAA A 1851

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 472 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Lys Phe Leu Ile Phe Ala Phe Phe Gly Gly Val His Leu Leu Ser
1 5 10 15

Leu Cys Ser Gly Lys Ala Ile Cys Lys Asn Gly Ile Ser Lys Arg Thr
20 25 30

Phe Glu Glu Ile Lys Glu Glu Ile Ala Ser Cys Gly Asp Val Ala Lys
35 40 45

Ala Ile Ile Asn Leu Ala Val Tyr Gly Lys Ala Gln Asn Arg Ser Tyr
50 55 60

Glu Arg Leu Ala Leu Leu Val Asp Thr Val Gly Pro Arg Leu Ser Gly
65 70 75 80

Ser Lys Asn Leu Glu Lys Ala Ile Gln Ile Met Tyr Gln Asn Leu Gln
85 90 95

Gln Asp Gly Leu Glu Lys Val His Leu Glu Pro Val Arg Ile Pro His
100 105 110

```
Trp Glu Arg Gly Glu Glu Ser Ala Val Met Leu Glu Pro Arg Ile His
        115                 120                 125

Lys Ile Ala Ile Leu Gly Leu Gly Ser Ser Ile Gly Thr Pro Pro Glu
    130                 135                 140

Gly Ile Thr Ala Glu Val Leu Val Val Thr Ser Phe Asp Glu Leu Gln
145                 150                 155                 160

Arg Arg Ala Ser Glu Ala Arg Gly Lys Ile Val Val Tyr Asn Gln Pro
                165                 170                 175

Tyr Ile Asn Tyr Ser Arg Thr Val Gln Tyr Arg Thr Gln Gly Ala Val
            180                 185                 190

Glu Ala Ala Lys Val Gly Ala Leu Ala Ser Leu Ile Arg Ser Val Ala
        195                 200                 205

Ser Phe Ser Ile Tyr Ser Pro His Thr Gly Ile Gln Glu Tyr Gln Asp
    210                 215                 220

Gly Val Pro Lys Ile Pro Thr Ala Cys Ile Thr Val Glu Asp Ala Glu
225                 230                 235                 240

Met Met Ser Arg Met Ala Ser His Gly Ile Lys Ile Val Ile Gln Leu
                245                 250                 255

Lys Met Gly Ala Lys Thr Tyr Pro Asp Thr Asp Ser Phe Asn Thr Val
            260                 265                 270

Ala Glu Ile Thr Gly Ser Lys Tyr Pro Glu Gln Val Val Leu Val Ser
        275                 280                 285

Gly His Leu Asp Ser Trp Asp Val Gly Gln Gly Ala Met Asp Asp Gly
    290                 295                 300

Gly Gly Ala Phe Ile Ser Trp Glu Ala Leu Ser Leu Ile Lys Asp Leu
305                 310                 315                 320

Gly Leu Arg Pro Lys Arg Thr Leu Arg Leu Val Leu Trp Thr Ala Glu
                325                 330                 335

Glu Gln Gly Gly Val Gly Ala Phe Gln Tyr Tyr Gln Leu His Lys Val
            340                 345                 350

Asn Ile Ser Asn Tyr Ser Leu Val Met Glu Ser Asp Ala Gly Thr Phe
        355                 360                 365

Leu Pro Thr Gly Leu Gln Phe Thr Gly Ser Glu Lys Ala Arg Ala Ile
    370                 375                 380

Met Glu Glu Val Met Ser Leu Leu Gln Pro Leu Asn Ile Thr Gln Val
385                 390                 395                 400

Leu Ser His Gly Glu Gly Thr Asp Ile Asn Phe Trp Ile Gln Ala Gly
```

405 410 415

Val Pro Gly Ala Ser Leu Leu Asp Asp Leu Tyr Lys Tyr Phe Phe Phe
420 425 430

His His Ser His Gly Asp Thr Met Thr Val Met Asp Pro Lys Gln Met
435 440 445

Asn Val Ala Ala Ala Val Trp Ala Val Val Ser Tyr Val Val Ala Asp
450 455 460

Met Glu Glu Met Leu Pro Arg Ser
465 470

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1076 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CAGAAGTTCA AGGGCCCCCG GCCTCCTGCG CTCCTGCCGC CGGGACCCTC GACCTCCTCA 60

GAGCAGCCGG CTGCCGCCCC GGGAAGATGG CGAGGAGGAG CCGCCACCGC CTCCTCCTGC 120

TGCTGCTGCG CTACCTGGTG GTCGCCCTGG GCTATCATAA GGCCTATGGG TTTTCTGCCC 180

CAAAAGACCA ACAAGTAGTC ACAGCAGTAG AGTACCAAGA GGCTATTTTA GCCTGCAAAA 240

CCCCAAAGAA GACTGTTTCC TCCAGATTAG AGTGGAAGAA ACTGGGTCGG AGTGTCTCCT 300

TTGTCTACTA TCAACAGACT CTTCAAGGTG ATTTTAAAAA TCGAGCTGAG ATGATAGATT 360

TCAATATCCG GATCAAAAAT GTGACAAGAA GTGATGCGGG GAAATATCGT TGTGAAGTTA 420

GTGCCCCATC TGAGCAAGGC CAAAACCTGG AAGAGGATAC AGTCACTCTG GAAGTATTAG 480

TGGCTCCAGC AGTTCCATCA TGTGAAGTAC CCTCTTCTGC TCTGAGTGGA ACTGTGGTAG 540

AGCTACGATG TCAAGACAAA GAAGGGAATC CAGCTCCTGA ATACACATGG TTTAAGGATG 600

GCATCCGTTT GCTAGAAAAT CCCAGACTTG GCTCCCAAAG CACCAACAGC TCATACACAA 660

TGAATACAAA AACTGGAACT CTGCAATTTA ATACTGTTTC CAAACTGGAC ACTGGAGAAT 720

ATTCCTGTGA AGCCCGCAAT TCTGTTGGAT ATCGCAGGTG TCCTGGGAAA CGAATGCAAG 780

TAGATGATCT CAACATAAGT GGCATCATAG CAGCCGTAGT AGTTGTGGCC TTAGTGATTT 840

CCGTTTGTGG CCTTGGTGTA TGCTATGCTC AGAGGAAAGG CTACTTTTCA AAAGAAACCT 900

CCTTCCAGAA GAGTAATTCT TCATCTAAAG CCACGACAAT GAGTGAAAAT GATTTCAAGC 960

ACACAAAATC CTTTATAATT TAAAGACTCC ACTTTAGAGA TACACCAAAG CCACCGTTGT 1020

TACACAAGTT ATTAAACTAT TATAAAACTC AAAAAAAAAA AAAAAAAAAA AAAAAA 1076

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 298 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Ala Arg Arg Ser Arg His Arg Leu Leu Leu Leu Leu Leu Arg Tyr
1 5 10 15

Leu Val Val Ala Leu Gly Tyr His Lys Ala Tyr Gly Phe Ser Ala Pro
20 25 30

Lys Asp Gln Gln Val Val Thr Ala Val Glu Tyr Gln Glu Ala Ile Leu
35 40 45

Ala Cys Lys Thr Pro Lys Lys Thr Val Ser Ser Arg Leu Glu Trp Lys
50 55 60

Lys Leu Gly Arg Ser Val Ser Phe Val Tyr Tyr Gln Gln Thr Leu Gln
65 70 75 80

Gly Asp Phe Lys Asn Arg Ala Glu Met Ile Asp Phe Asn Ile Arg Ile
85 90 95

Lys Asn Val Thr Arg Ser Asp Ala Gly Lys Tyr Arg Cys Glu Val Ser
100 105 110

Ala Pro Ser Glu Gln Gly Gln Asn Leu Glu Glu Asp Thr Val Thr Leu
115 120 125

Glu Val Leu Val Ala Pro Ala Val Pro Ser Cys Glu Val Pro Ser Ser
130 135 140

Ala Leu Ser Gly Thr Val Val Glu Leu Arg Cys Gln Asp Lys Glu Gly
145 150 155 160

Asn Pro Ala Pro Glu Tyr Thr Trp Phe Lys Asp Gly Ile Arg Leu Leu
165 170 175

Glu Asn Pro Arg Leu Gly Ser Gln Ser Thr Asn Ser Ser Tyr Thr Met
180 185 190

Asn Thr Lys Thr Gly Thr Leu Gln Phe Asn Thr Val Ser Lys Leu Asp
195 200 205

Thr Gly Glu Tyr Ser Cys Glu Ala Arg Asn Ser Val Gly Tyr Arg Arg
210 215 220

Cys Pro Gly Lys Arg Met Gln Val Asp Asp Leu Asn Ile Ser Gly Ile
225 230 235 240

Ile Ala Ala Val Val Val Val Ala Leu Val Ile Ser Val Cys Gly Leu
245 250 255

Gly Val Cys Tyr Ala Gln Arg Lys Gly Tyr Phe Ser Lys Glu Thr Ser
260 265 270

Phe Gln Lys Ser Asn Ser Ser Ser Lys Ala Thr Thr Met Ser Glu Asn
275 280 285

Asp Phe Lys His Thr Lys Ser Phe Ile Ile
290 295

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2522 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAAAGGTTGT GTAGCTTGCC CTGGTTGCAT AGTTAAACGA GGGCTAGAAA CAGGACTAGG 60

AGTCAGGCCT GTCCAGCTGG AAAACTTGGG TTTTCTAGAA GGGGTACCCT GGCCTCCTGC 120

GGAGCCTGCT GTGGGACTCT GCAGAACACA ATTCAAGGCC AGACTGAACA CTAGCCTGAA 180

CCTGCCCTGA GAATCCCTCT AAGCCGACCT ACTCCACAGC TGTCCTGACT GTGTAAGCGA 240

GATGATGATT AGTGATCAGA CGAAAGGATT CCTGTCATTG GTAACCCTCT CAAAGTATTT 300

GGAAAACAGT TCAATTTTCA TCTATTTCAG AAGCACGCCG TGGTGTCTAT TGAGGCTCAC 360

CTGCATTGAA TTCCTTCCTT TTTATGTTGC GATCTCCCAA GATTGCATTG TGGAGTGTTT 420

TCGAATCCAT TTTGAAATCC CCGTGCGTGC GCTATGCAGG CCTCAGTCTT TTTCCATTCC 480

ATTCTTAACT CTACTTTCGA CGGAAGCAGT GTTTTACCCC GACACTGGCT TGCCTAGGAC 540

CTTGTGCTCT GCACAACTAG CAGGGCCCGG CAGGATGTAC TGAATTCTTG CTCTCGTGTC 600

CAGCTGGACG GTGATGGCTT TCAAGTCCTT GGCTGTTGGG AGCTTACTAT AAATGTTCGT 660

CTTGGCTACA AACTCTCCAC TCTTTCCTCG GCACTCTCTC AGCATTGCCA CCACTGTCTT 720

TCCTCTTGGC CAACTGTTTT CTTTACTTAG GCTTTCCCTT GCTAGAAAGT CCAGGTAACT 780

TTCTCCACGG GACCTGGTTT CCTTCGCACA TCCCAGCTGG CCTCGAGGAA AGGTAGCTCT 840

TCCAAATCAG AGAATCTGGA TGCTGGGCTG GGCTCTGCAC CAACCAGCTG GGCCGCTTCA 900

CCCGCTGGGC CCCAAACTAC TCATCTGTGA AGCGAAGGCA CCGCGCTTGA TGCCTTCTGC 960

AACGTTCTTC AGTTTGGAAA TCCTTCTGTT TCGTTGGGGA TATTTCACGG CCTCTTCTCA 1020

AGGTTGCACT TTTGCCAGCT GCCAGGGATC GTCTCAAAAC AGGTTCTTAG TGCATTCATA 1080

GCTTGAGCTG CTGTCTTGAA AGTAGTACAT TCCTTTTTCT GCCAACTTTT TTCTGAGAAA 1140

GTTTTTGAAT GCACACGTGC ACCCAACAGA GTGAGAGTGG CTGTTAAGAG AGAGGGCGCC 1200

ATTTCCTTTG CCCTCCAGCC TGTCCCTGTG CACCCTGGAG GGGCCCGTTT TTTCCACCGC 1260

TTAGATAAAA TCTAGGGCAA GTTCCTGAAC TTCTCTTGTC TCTCTCAGGT AACAAAAATT 1320

CTTTTGGGCT CCTTTAGTCA CAAAGATATT CACGATTTCA GGTATTAAAG TGCCCAGCCC 1380

TGGGTGATTG TCAAAATTCT GAACTTGATT TAAAGTGGCA CCTCCTCTCA CAGTCTTCGG 1440

GAGGGAGAGA CCGGAGCCAG GAGTGCAGCG TGTTTGCTGG GGTCTGTCGT GGCCCACTCC 1500

ACACCTGCTG GGTGGATCCG GCTGGTGCCC CATGGGCGCC TCTGAGATGC CCCTCCCCAC 1560

CCCATCAGTG GCGCTGTCTC ACCTGCAGGC TGTTCTCACA GGTGGTCCCC CCTCACTCCT 1620

CCTGCAGCCC CAGTTCCTGG CTGTTCATTC TTATTGGGAC CCGTCACCCT CCTGGAGGCG 1680

GTCCCAGCCG AGCCCCCTTA AGACAGCACC AGGCTGGCTC CACTTGGCCC CCGCTGGTTC 1740

AGGGAAGTGC TGCTGCAGCC GTTTAGTTTG ACAAAGGAGG CAGCGAGGCC GTCTCATTGG 1800

TAGCCCTCTC CTGGCTTGCC CAGCCACCAC CTCACCTCGA TTCCTCCCAG GCCTGGGTCC 1860

AGCACCAGCC TAGGAAGAGG GTGCCCCATG CTGTCTAGCT CTTCTTCGGG ATGGGGGGCT 1920

CCAGGTTCCT TGGTATTTTG CTTTGGCCTT TGGAGCCTCA GTCAAAACTG AGGAAAGGTG 1980

TCATTTTCAC ATCTCGTCAC ACGTACAGTG ACTGCAACTA AAAGCACAGG CTTTGTAGAA 2040

ACAGACATGG GTTCAGGCCC CAGCTCCACC ATTCACAAGG TGTGTGGCTT CCTGCAAGGT 2100

ACCTTCATCT CTGAGTTACC TGACTCCATC TGAGTTTCCT TCTTGTAAAA CTGGCATCCA 2160

TGAAAGTGGC TACCTCGAAG GGCGTGAAGA TGAAATGAGG TGGAAAGTAG GTAGCCCCCG 2220

AATGAGGGAA GCATTGAGTG AGAGCTGGCC CTCTGACCCT TCTAAAAGAA CACAGCCAAC 2280

TTTTTAAACT GTCTTTCCAG AAAGAGATGG AAAACTTCGA AGCCCCTTTC CACTGCCTTG 2340

CCAAGCAGTT CCACCAGCTG TACCGGGAGA AGGTGGAGGT TTTCCGGGCC CTGGCATGAC 2400

GAGCTGGAGC AGATCGTGCT GCACAACCGG AGAAGACAGA ATTACCTCTG CTCTTTTAAT 2460

ATATAATGAT GGCTTTAAAT AAAATTAGGA GAAAATGTCA AAAAAAAAAA AAAAAAAAAA 2520

AA 2522

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 113 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Met Ile Ser Asp Gln Thr Lys Gly Phe Leu Ser Leu Val Thr Leu
1 5 10 15

Ser Lys Tyr Leu Glu Asn Ser Ser Ile Phe Ile Tyr Phe Arg Ser Thr
20 25 30

Pro Trp Cys Leu Leu Arg Leu Thr Cys Ile Glu Phe Leu Pro Phe Tyr
35 40 45

Val Ala Ile Ser Gln Asp Cys Ile Val Glu Cys Phe Arg Ile His Phe
50 55 60

Glu Ile Pro Val Arg Ala Leu Cys Arg Pro Gln Ser Phe Ser Ile Pro
65 70 75 80

Phe Leu Thr Leu Leu Ser Thr Glu Ala Val Phe Tyr Pro Asp Thr Gly
85 90 95

Leu Pro Arg Thr Leu Cys Ser Ala Gln Leu Ala Gly Pro Gly Arg Met
100 105 110

Tyr

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1962 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCCGGGCCCC AGCCTTCTCC AGAACCCCTG CTACCCACGA CTAAGCCCCG AACAATCTGC 60

CCTTGGGCTT GTTCTCTTCG CAGTTGTCGG CCCTGGGCCG GGAGCTGGAG TCCCAGACTC 120

ATAGGTCCCG GCCCAGCCCC CGAAGAGCCG CCTCAGCCGG GGGGAGTTGC TCGGACTCAA 180

ACGTCCAGTC CTCGTGCGAC CGCGCTGGGT CGGAAGTGAG CAGGCTGAGG CCACCATGGA 240

GCAGTGTGCG TGCGTGGAGA GAGAGCTGGA CAAGGTCCTG CAGAAGTTCC TGACCTACGG 300

GCAGCACTGT GAGCGGAGCC TGGAGGAGCT GCTGCACTAC GTGGGCCAGC TGCGGGCTGA 360

GCTGGCCAGC GCAGCCCTCC ARGGGACCCC TCTCTCAGCC ACCCTCTCTC TGGTGATGTC 420

ACAGTGCTGC CGGAAGATCA AAGATACGGT GCAGAAACTG GCTTCGGAMC ATAAGGACAT 480

TCACAGCAGT GTATCCCGAG TGGGCAAAGC CATTGACAGG AACTTCGACT CTGAGATCTG 540

TGGTGTTGTG TCAGATGCGG TGTGGGACGC GCGGGAACAG CAGCAGCAGA TCCTGCAGAT 600

GGCCATCGTG GAACACCTGT ATCAGCAGGG CATGCTCAGC GTGGCCGAGG AGCTGTGCCA 660

GGAATCAACG CTGAATGTGG ACTTGGATTT CAAGCAGCCT TTCCTAGAGT TGAATCGAAT 720

CCTGGAAGCC CTGCACGAAC AAGACCTGGG TCCTGCGTTG GAATGGGCCG TCTCCCACAG 780

GCAGCGCCTG CTGGAACTCA ACAGCTCCCT GGAGTTCAAG CTGCACCGAC TGCACTTCAT 840

CCGCCTCTTG GCAGGAGGCC CCGCGAAGCA GCTGGAGGCC CTCAGCTATG CTCGGCACTT 900

CCAGCCCTTT GCTCGGCTGC ACCAGCGGGA GATCCAGGTG ATGATGGGCA GCCTGGTGTA 960

CCTGCGGCTG GGCTTGGAGA AGTCACCCTA CTGCCACCTG CTGGACAGCA GCCACTGGGC 1020

AGAGATCTGT GAGACCTTTA CCCGGGACGC CTGTTCCCTG CTGGGGCTTT CTGTGGAGTC 1080

CCCCCTTAGC GTCAGCTTTG CCTYTGGCTG TGTGGCGCTG CCTGTGTTGA TGAACATCAA 1140

GGCTGTGATT GAGCAGCGGC AGTGCACTGG GGTCTGGAAT CACAAGGACG AGTTACCGAT 1200

GAGATTGAAC TAGGCATGAA GTGCTGGTAC CACTCCGTGT TCGCTTGCCC CATCCTCCGC 1260

CAGCAGACGT CAGATTCCAA CCCTCCCATC AAGCTCATCT GTGGCCATGT TATCTCCCGA 1320

GATGCACTCA ATAAGCTCAT TAATGGAGGA AAGCTGAAGT GTCCCTACTG TCCCATGGAG 1380

CAGAACCCGG CAGATGGGAA ACGCATCATA TTCTGATTCC TACCTGGAAG GAATTTTGTT 1440

GAAAGGGGTT TTCACCTGTG AGCCTTGGTC TGTCTCGGTA GGGTGGTCAA CTTCAGTGGA 1500

CTGTGGTTGG TTTCAGAGCG CCTGGCTGAG GAGTTCCACT GAGGGGAGCA CTGGAGCAGC 1560

CCTTTGGCAG AGGCTGAGGA GGGAGATGGA CCAGCCCACG CCTGGCACCT GGCTCCATGG 1620

CATAAGGAAA GGGAGATGCT GGCCTCTGTG CTCCTGCTGT CTTTTCCTGT TTCTGTTTGC 1680

GTTTGACTTA GTAGCAACCG ACAGAGTGGC AAGGGATTTG GTCTTCAGCA GTAGACATCC 1740

TTCCACCCCT GCCCTCAGCC AAGTCTCTTG CTGCCATGCC AATGCTATGT CCACCCTTGC 1800

CCCTCGGCCC AAGAGTGTCC AGCGGTGGCC CACYTYTTCC TCCCACTACA GCCTCAACAG 1860

TATGTACCAT CTCCCACTGT AAATAGTCCC AGTTAGAACG GAATGCCGTT GTTTTATAAC 1920

TTTGAACAAA TGTATTTACT GCCAAAAAAA AAAAAAAAAA AA 1962

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 325 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Glu Gln Cys Ala Cys Val Glu Arg Glu Leu Asp Lys Val Leu Gln
1 5 10 15

Lys Phe Leu Thr Tyr Gly Gln His Cys Glu Arg Ser Leu Glu Glu Leu
20 25 30

Leu His Tyr Val Gly Gln Leu Arg Ala Glu Leu Ala Ser Ala Ala Leu
35 40 45

Gln Gly Thr Pro Leu Ser Ala Thr Leu Ser Leu Val Met Ser Gln Cys
50 55 60

Cys Arg Lys Ile Lys Asp Thr Val Gln Lys Leu Ala Ser Xaa His Lys
65 70 75 80

```
Asp Ile His Ser Ser Val Ser Arg Val Gly Lys Ala Ile Asp Arg Asn
                85                  90                  95

Phe Asp Ser Glu Ile Cys Gly Val Val Ser Asp Ala Val Trp Asp Ala
            100                 105                 110

Arg Glu Gln Gln Gln Gln Ile Leu Gln Met Ala Ile Val Glu His Leu
        115                 120                 125

Tyr Gln Gln Gly Met Leu Ser Val Ala Glu Glu Leu Cys Gln Glu Ser
    130                 135                 140

Thr Leu Asn Val Asp Leu Asp Phe Lys Gln Pro Phe Leu Glu Leu Asn
145                 150                 155                 160

Arg Ile Leu Glu Ala Leu His Glu Gln Asp Leu Gly Pro Ala Leu Glu
                165                 170                 175

Trp Ala Val Ser His Arg Gln Arg Leu Leu Glu Leu Asn Ser Ser Leu
            180                 185                 190

Glu Phe Lys Leu His Arg Leu His Phe Ile Arg Leu Leu Ala Gly Gly
        195                 200                 205

Pro Ala Lys Gln Leu Glu Ala Leu Ser Tyr Ala Arg His Phe Gln Pro
    210                 215                 220

Phe Ala Arg Leu His Gln Arg Glu Ile Gln Val Met Met Gly Ser Leu
225                 230                 235                 240

Val Tyr Leu Arg Leu Gly Leu Glu Lys Ser Pro Tyr Cys His Leu Leu
                245                 250                 255

Asp Ser Ser His Trp Ala Glu Ile Cys Glu Thr Phe Thr Arg Asp Ala
            260                 265                 270

Cys Ser Leu Leu Gly Leu Ser Val Glu Ser Pro Leu Ser Val Ser Phe
        275                 280                 285

Ala Xaa Gly Cys Val Ala Leu Pro Val Leu Met Asn Ile Lys Ala Val
    290                 295                 300

Ile Glu Gln Arg Gln Cys Thr Gly Val Trp Asn His Lys Asp Glu Leu
305                 310                 315                 320

Pro Met Arg Leu Asn
                325
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 745 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AAAAACACAA AACCCCGTAA AATCACAAAG AAAATCCAAC ACCAAAGGCG CAGAAGCCGG 60

CTGGCCGTGG TGGGGGCAGC GTAGGCGTAG CATCCCTCTC CTCTCACTTA GCCTGTTGAC 120

TCTTGTTATT ATCATGATAT TCACAAAACG CCGCATGTTT AAAAAGTCAT AGATGTCATC 180

TTCTCTCTGC CCCCAGGGAG GAAAGCCACC TTCTCTTGCC CCTTGGCCCC TTTGTCAGGG 240

GCCAGGGGTC TGCCGGGTGG GGGTGCCAAC AGGCCTGGCC CTTTCCTCCC CTGCATCCAG 300

CCATGGGGGC CTCTGCGATT GCCGGAAGGT TGCATGGCTG GTCCCAGGGC CAGCACAGGC 360

CCGAGGCCGG GCTGCCTGGT TTTATTTTTA TTTAACTTTA TTTTCTGTTT TATGAGTGTG 420

TGTCCGCCCA CCCCCACCCC CTTCAGTGTT AAGTGGGGAG CCCTGGGGGA GTCTCTCCTG 480

CCTCCCAGCC TCTCCCAAGA CCTCCCCCCT CGTCACCAGC CATCCCTCTG GACCAGGCAG 540

AGGGCGGACC GGGTGGGCAG GGGCCTGAGG GTGGCTCGGG CCAGCCCACC AGCCAATGGA 600

CCCCTCCTCA GGCCGCCAGT GTCGCCCTGC CCCTTTTTAA AACAAAATGC CCTCGTTTGT 660

AAACCCTTAG ACGCTTGAGA ATAAACCCCT TCCTTTTCTT CCAAAAAAAA AAAAAAAAAA 720

AAAAAAAAAA AAAAAAAAAA AAAAA 745

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 114 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Ala Gly Pro Arg Ala Ser Thr Gly Pro Arg Pro Gly Cys Leu Val
1 5 10 15

Leu Phe Leu Phe Asn Phe Ile Phe Cys Phe Met Ser Val Cys Pro Pro
20 25 30

Thr Pro Thr Pro Phe Ser Val Lys Trp Gly Ala Leu Gly Glu Ser Leu

35 40 45

Leu Pro Pro Ser Leu Ser Gln Asp Leu Pro Pro Arg His Gln Pro Ser
50 55 60

Leu Trp Thr Arg Gln Arg Ala Asp Arg Val Gly Arg Gly Leu Arg Val
65 70 75 80

Ala Arg Ala Ser Pro Pro Ala Asn Gly Pro Leu Leu Arg Pro Pro Val
85 90 95

Ser Pro Cys Pro Phe Leu Lys Gln Asn Ala Leu Val Cys Lys Pro Leu
100 105 110

Asp Ala

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1983 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TGGCAATAGT GGTTAGGGAA GGCTCCTTTG AGGAAGTGAA TTTTTAGCTG AGACTTAAAG 60

AACAAATGAG ATTTAGCTAG AAAAATTGGA CATGCGATGC CAAGATGGCA TTTTAAAAGA 120

ATAATAGTAA GCACAAAGGC CCTGTAGCAG GAGGGAGCTG ATTGTCCATA GTTCAGACAG 180

CAAAGAAGCT GATGATGCAG GTTGGGGTCA GACCGTGTTT GACTACAGAT AGGATGTTAA 240

GGGTTTTGGC TTTTTAGGTT TTTGTTTTAA TTCTAAAAGT AATGGAAAAT GTACTCCTTT 300

TGGTGGTGGT CTGAGAGAAG GTACATCATT AGAATGACAT TTTGAAAACA ACACTCAGGC 360

TGCTCAGTAG AGAATGGCTT CAAAGGATTT AAAAGCAGAA GCAGAAGGAC ATATTAGAGA 420

AGGATTGTAT AGTTTTCTGG TAAAAGATGA CAGTGAATTG TATGGGCGAT GGATTAGCCG 480

TGGAAGGTGT TGAGTATAAG TGGTCTCCAG CCAAACTCTA TGGTTACTGG AATAAGAGAG 540

TAGGAACCCT TCTCAGGCTT TATCTTTATC TATTCTTGTC AACAGTATGT ACATGTGTCC 600

CCCAGCCCCA AATAACTGTA CAGTTTAATG ATGTTCACTC TATACAGTTC CCAGAATCCA 660

TTGGAAATTG CTGTAACAGC ATATCCTCAA TGCCCATCAA TTCTCCACGT CCAACTTCTC 720

CATGGCCTCC TCTGCCTCTG CTGATCTGTG AACTTCCCAA GCCCCTTCCC CTACCTGCTT 780

TTGATTGGCT TTAACTTTTA CAATATCTTC ATTACTCCAA GTTTGTTCAA CATCCTTTTT 840

ATTTTTTTAA ATCATAGATT GATTTAGTTT ATTCTCTTTG CCATTTTTGA ATCTCATTAT 900

TTCTGTTTCT CCTTGGTTAT TAGTGGCTCT GTTTTCCTTC AATTGCCTCT TGTCTTTGAG 960

AAGCTCTTGT GATTCTTTTA GGGCCATTTG CCATTTGATT GGTTTGTCTT CCTTTTCCCT 1020

ATAAGCTTTA AATATGGCAT TATAGTTTTA TCCCCTTTCC TCTTCTTTAG GTACAACTGC 1080

AGACACTTTG CTCTTCCAAG GTTACTAAGC AGTGTCTGAC ACAATGTAGA AGCTCAACAA 1140

ATATTGGTTA AATTTATTTC TTCTATTGAT TGTTCAGGCT TTGATGACAT CACTTAAAAT 1200

GTTTCTTGTA CACACCCTGT TTTCTACTGA TATATGTATG TGTATGGCTA CCTGAATCCA 1260

GGTTTCTTCT AGGAATATAC AGAAAGTAAT TGATTTCTCT GTGGATCTCT AACAGTGACA 1320

AGAATTTTCA CCTATGCCTG TGAGAATACC TTCAAAAGTA TTGGGTGCTC ATCATAAACA 1380

CACATCAGTT TAACAAACTC TTATGGATGC ATTGACTTTC CCAGTTAGTT GCTAGATGAC 1440

TTCGGATGAT TTGCATAATG GGTCTCAGTT TCCATATCTG TTAAATGGCA ATAATCAGAG 1500

AATTTTAAAA AATTTAAGGA CACCTGGAAA GCTTGAAAGA TCCCTAGAAA GCATGTGTTT 1560

ATTCCACATA GTGGGAACTA TGCTAGATTC CCAAAGACAC AAAGACAACT AAGACAACTT 1620

AGAATAAGAA GGAAAAGAGA ATGATTCGTT GCAATGATCC CCTTGAAGCT CCAGTTGAAA 1680

GTCAGAGTAT TGCCCTGGAT TGGAAGTAGT CTCCAAACTG ACATCATTTT CTTTTTCGAA 1740

CCATATCTGG CCTGTCTCTC TTGCCAGTTG CATATTAAAG GTAACAGATT TGAAAATGTT 1800

TGGAATAAAA GCTCTAGTTA GGTGTGGTGG CACACACCTG CCATCCCAGC TACTGGAGAG 1860

TCTGAGACTC GATGATTGTT TCAGCCCAAG AGTTGGAGGT TGTAGTGAGC TATGATGGCA 1920

CCACTGCACT CCAGTCTGTG TGACAGAGCG AAGACCTTGT CTCTAAGGAA AAAAAAAAAA 1980

AAA 1983

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 115 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Thr Val Asn Cys Met Gly Asp Gly Leu Ala Val Glu Gly Val Glu
1 5 10 15

Tyr Lys Trp Ser Pro Ala Lys Leu Tyr Gly Tyr Trp Asn Lys Arg Val
20 25 30

Gly Thr Leu Leu Arg Leu Tyr Leu Tyr Leu Phe Leu Ser Thr Val Cys
35 40 45

Thr Cys Val Pro Gln Pro Gln Ile Thr Val Gln Phe Asn Asp Val His
50 55 60

Ser Ile Gln Phe Pro Glu Ser Ile Gly Asn Cys Cys Asn Ser Ile Ser
65 70 75 80

Ser Met Pro Ile Asn Ser Pro Arg Pro Thr Ser Pro Trp Pro Pro Leu
85 90 95

Pro Leu Leu Ile Cys Glu Leu Pro Lys Pro Leu Pro Leu Pro Ala Phe
100 105 110

Asp Trp Leu
115

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1046 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GGGCTTAGTT AGGAGCTATG GCTAAACATC ATCCTGATTT GATCTTTTGC CGCAAGCAGG 60

CTGGTGTTGC CATCGGAAGA CTGTGTGAAA AATGTGATGG CAAGTGTGTG ATTTGTGACT 120

CCTATGTGCG TCCCTGCACT CTGGTGCGCA TATGTGATGA GTGTAACTAT GGATCTTACC 180

AGGGGCGCTG TGTGATCTGT GGAGGACCTG GGGTCTCTGA TGCCTATTAT TGTAAGGAGT 240

GCACCATCCA GGAGAAGGAC AGAGATGGCT GCCCAAAGAT TGTCAATCTG GGGAGCTCTA 300

AGACAGACCT CTTCTATGAA CGCAAAAAAT ACGGCTTCAA GAAGARGTGA TTGGTGGGTG 360

GCCCCTTCCT CCCCCCAACA TCAGTCTGCT GCAGCTGCCA GAAAACATGC CTACTACTAC 420

CAGCAGAAAG GGAGCAGAGC CCAGAGCATC ACCAGGAGTG CCTGCTAGTG TACTGGCAGC 480

TTGCCACCCC CTCCTCTCCC TTCACCCAGA CACGTGGTAG GGATGGAAAA GGATTCTTCA 540

CAGAGCACTC TGGCACACCA TATCGGAGAA AACTTGATAG ATTAGTTAAT GGTTTTTCTT 600

GAATTCGAGA AGCATAGATC TGTTCTCCAT ATTGGTATGT TCTCCCTCAA CCAAGATCTT 660

CTAAAAAGAA ATAATATTTT AGTCTTCTGC TTGAGGAACT GACTGTGAAG CGACGCCCAG 720

TGAAAAACAT GTTCTTGCAG CAGCTCTGGT GGCAGCTGTC CTTGAGGAAC CTTTGGTGTG 780

TGGTGGGAAG CTATCAGAAC AAGAAATGTA GGCATTTCCC GTTTTTTTGG GGGGGGGGGG 840

TGGGGGGGCA GGGCTCTGCC CTCTTGAAAG GCATTTACTT GTTTAACACT TGTCCAGCTA 900

CAGTGGGGTA CAGTAGCTGG CTATTCACAG GCATCATCAT AGCCCACTAG TCTCATATTA 960

TTTTCCTTTT GAGAAATTGG AAACTCTTTC TGTTGCTATT ATATTAATAA AGTTGGTGTT 1020

TATTTTCTGG TAAAAAAAAA AAAAAA 1046

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 110 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Ala Lys His His Pro Asp Leu Ile Phe Cys Arg Lys Gln Ala Gly
1 5 10 15

Val Ala Ile Gly Arg Leu Cys Glu Lys Cys Asp Gly Lys Cys Val Ile
20 25 30

Cys Asp Ser Tyr Val Arg Pro Cys Thr Leu Val Arg Ile Cys Asp Glu
35 40 45

Cys Asn Tyr Gly Ser Tyr Gln Gly Arg Cys Val Ile Cys Gly Gly Pro
50 55 60

Gly Val Ser Asp Ala Tyr Tyr Cys Lys Glu Cys Thr Ile Gln Glu Lys
65 70 75 80

Asp Arg Asp Gly Cys Pro Lys Ile Val Asn Leu Gly Ser Ser Lys Thr
85 90 95

Asp Leu Phe Tyr Glu Arg Lys Lys Tyr Gly Phe Lys Lys Xaa
100 105 110

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1814 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ATCTTGCAGT GGGCCTCTGT CCCAAAAACA AGCAGAATTT TTTCTTTCTC AACAGGCTTC 60

TTTGCTAAAG AATGATGAGA CTAAGGCCCT CACTCCAGCT TCCTTGCAGA AGGAATTAAA 120

CAATTTGTTG AAATTTAATC CTGATTTTGC TGAAGCGCAT TATCTCAGCT ACTTAAACAA 180

CCTCCGTGTC CAAGATGTTT TCAGTTCAAC ACACAGTCTC CTCCATTATT TTGATCGTCT 240

GATTCTTACC GGAGCCGAAA GCAAAAGTAA TGGGGAAGAR GGCTATGGCC GGAGCTTGAG 300

ATACGCCGCT CTGAATCTTG CCGCCCTGCA CTGCCGCTTC GGTCACTATC AACAGGCAGA 360

GCTCGCCCTG CAGGARGCAA TTAGGATTGC CCAGGARTCC AACGATCACG TGTGTCTCCA 420

GCACTGTTTG AGCTGGCTTT ATGTGCTGGG GCAGAAGAGA TCCGATAGCT ATGTTCTGCT 480

GGAGCATTCT GTGAAGAAGG CAGTACATTT TGGGTTACCG TACCTCGCCT CCCTGGGAAT 540

ACAGTCCCTT GTTCAACAGA GAGCTTTTGC TGGGAAGACG GCAAACAAGC TGATGGATGC 600

CCTAAAGGAC TCCGACYTCC TGCACTGGAA ACACAGCCTG TCAGAGCTCA TCGATATCAG 660

CATCGCACAG AAAACGGCCA TCTGGAGGCT GTATGGCCGC AGCACCATGG CACTGCAACA 720

GGCCCAGATG TTGCTGAGCA TGAACAGCCT GGAGGCGGTG AATGCGGGCG TGCAGCAGAA 780

CAACACAGAG TCCTTTGCTG TCGCACTCTG CCACCTCGCA GAGCTACACG CGGAGCAGGG 840

CTGTTTTGCT GCAGCTTCTG AAGTGTTAAA GCACTTGAAG GAACGATTTC CGCCTAATAG 900

TCAGCACGCC CAGTTATGGA TGCTATGTGA TCAAAAAATA CAGTTTGACA GAGCAATGAA 960

TGATGGCAAA TATCATTTGG CTGATTCACT TGTTACAGGA ATCACAGCTC TCAATAGCAT 1020

AGAGGGTGTT TATAGGAAAG CGGTTGTATT ACAAGCTCAG AACCAAATGT CAGAGGCACA 1080

TAAGCTTTTA CAAAAATTGT TGGTTCATTG TCAGAAACTG AAGAACACAG AAATGGTGAT 1140

CAGTGTCCTA CTGTCCGTGG CAGAGCTGTA CTGGCGATCT TCCTCCCCTA CCATCGCGCT 1200

GCCCATGCTC CTGCAGGCTC TGGCCCTCTC CAAGGAGTAC CGGTTACAGT ACTTGGCCTC 1260

TGAAACAGTG CTGAACTTGG CTTTTGCGCA GCTCATTCTT GGAATCCCAG AACAGGCCTT 1320

AAGTCTTCTC CACATGGCCA TCGAGCCCAT CTTGGCTGAC GGGGCTATCC TGGACAAAGG 1380

TCGTGCCATG TTCTTAGTGG CCAAGTGCCA GGTGGCTTCA GCAGCTTCCT ACGATCAGCC 1440

GAAGAAAGCA GAAGCTCTGG AGGCTGCCAT CGAGAACCTC AATGAAGCCA AGAACTATTT 1500

TGCAAAGGTT GACTGCAAAG AGCGCATCAG GGACGTCGTT TACTTCCAGG CCAGACTCTA 1560

CCATACCCTG GGGAAGACCC AGGAGAGGAA CCGGTGTGCG ATGCTCTTCC GGCAGCTGCA 1620

TCAGGAGCTG CCCTCTCATG GGGTACCCTT GATAAACCAT CTCTAGAGAG GACATCCCTG 1680

CTGGGCTGCT GTGCAGAGTA TAAGATTTTG GACTTGTTCA TGTCCCCTCT CTCCCTATAA 1740

ATGATGTATT TGTGACACCC TATCTTGTCA ATAAACAGCA TTCTGATTAG TTTGTCTTAA 1800

AAAAAAAAAA AAAA 1814

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 357 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Asp Ala Leu Lys Asp Ser Asp Xaa Leu His Trp Lys His Ser Leu
1 5 10 15

Ser Glu Leu Ile Asp Ile Ser Ile Ala Gln Lys Thr Ala Ile Trp Arg
20 25 30

Leu Tyr Gly Arg Ser Thr Met Ala Leu Gln Gln Ala Gln Met Leu Leu
35 40 45

Ser Met Asn Ser Leu Glu Ala Val Asn Ala Gly Val Gln Gln Asn Asn
50 55 60

Thr Glu Ser Phe Ala Val Ala Leu Cys His Leu Ala Glu Leu His Ala
65 70 75 80

Glu Gln Gly Cys Phe Ala Ala Ala Ser Glu Val Leu Lys His Leu Lys

```
                85                  90                  95

Glu Arg Phe Pro Pro Asn Ser Gln His Ala Gln Leu Trp Met Leu Cys
            100                 105                 110

Asp Gln Lys Ile Gln Phe Asp Arg Ala Met Asn Asp Gly Lys Tyr His
        115                 120                 125

Leu Ala Asp Ser Leu Val Thr Gly Ile Thr Ala Leu Asn Ser Ile Glu
    130                 135                 140

Gly Val Tyr Arg Lys Ala Val Val Leu Gln Ala Gln Asn Gln Met Ser
145                 150                 155                 160

Glu Ala His Lys Leu Leu Gln Lys Leu Leu Val His Cys Gln Lys Leu
                165                 170                 175

Lys Asn Thr Glu Met Val Ile Ser Val Leu Leu Ser Val Ala Glu Leu
            180                 185                 190

Tyr Trp Arg Ser Ser Ser Pro Thr Ile Ala Leu Pro Met Leu Leu Gln
        195                 200                 205

Ala Leu Ala Leu Ser Lys Glu Tyr Arg Leu Gln Tyr Leu Ala Ser Glu
    210                 215                 220

Thr Val Leu Asn Leu Ala Phe Ala Gln Leu Ile Leu Gly Ile Pro Glu
225                 230                 235                 240

Gln Ala Leu Ser Leu Leu His Met Ala Ile Glu Pro Ile Leu Ala Asp
                245                 250                 255

Gly Ala Ile Leu Asp Lys Gly Arg Ala Met Phe Leu Val Ala Lys Cys
            260                 265                 270

Gln Val Ala Ser Ala Ala Ser Tyr Asp Gln Pro Lys Lys Ala Glu Ala
        275                 280                 285

Leu Glu Ala Ala Ile Glu Asn Leu Asn Glu Ala Lys Asn Tyr Phe Ala
    290                 295                 300

Lys Val Asp Cys Lys Glu Arg Ile Arg Asp Val Val Tyr Phe Gln Ala
305                 310                 315                 320

Arg Leu Tyr His Thr Leu Gly Lys Thr Gln Glu Arg Asn Arg Cys Ala
                325                 330                 335

Met Leu Phe Arg Gln Leu His Gln Glu Leu Pro Ser His Gly Val Pro
            340                 345                 350

Leu Ile Asn His Leu
        355
```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1540 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CAGAATGTCT TAACATGAGA ATTGAATTTC ATGATGTTTG GTTCCATTTA ATAGCGGACA 60

CCACCCCAAT CTCATGTTTT CCTGTTACCC TAAAACAGTG GAAGGAAACT GGGTGTTTGG 120

TAGACTTCTA AATCATGGTC TCTGACAATT TGAATCTGAG ATTCTCACCT CCATTTACTA 180

AAGAATCGTG ACTTAATTCA AATTGCACAG TAATCAGTAA AGTGAATACG TTTTTAAAAT 240

GGAATTTTCT CCCTTCAGCA AGCACTCATT AAGGAGTGAG GCTGAGTATT TTAAGATAGA 300

GTGAGATCTG TGAGTGATTG AAAGGTGATA TTTAAAAACT TGGATTTCAT TCCAGTGTCA 360

GGTTTGGGTT TTAAGTTCCT TTGGTCCAGG GAAGGGTCCA AGCAGCCACA GTTGCCCTAA 420

ATCTCCATCA TTAAGTCTTC CAGCAAGGTT AAGTGCAGTA TGGAAGGAGA AGGGGGAAGA 480

GGACGGTAAC GGCCCCACAC TCCAGGCTGA GAAAGAGTAA TTAGGAGGCC TGAGGAGGGG 540

CCGAGGAAAG GCTGTTGGGG TGTGCTGGGG TTGGTACCCG AGCGCCTTCC CCTCACCTCA 600

ACCAGAGAAG AGCATCCGGT TGCTTTTTAA AGCTTTTAGC CTGCCCTAGC AAGGACAAAG 660

CATGTTAGAT TAGAGATGCT TCTGCTGATC GCAGGGGTTC TTATTTGAAA ACATCTATGA 720

TGGGGGTGGG GTGGGAGGAG ACAGGTTGTG GTTATGCAGG AAAATCTTGT CCTAAAAATA 780

TATGAGTTTG GGGGTAAGGG GTGGGATAGC CAAGCAAAAT CAGTAATTAT TTTAAAATGA 840

ACATATGAAT TTTTATTAAC TTTTAGTTAA ATACAGATTT TACAACGAGG TCAGCATAAG 900

CCTAAATCTA TATAGAGGGC TAACTCAGGC ATTGTCTTGT TTATTTGTAG ACTGGATTAA 960

AAACAACCTG TCCTGTTTTG TCAGTTCCCA GCTTCTTCGT TTAGAATAAA TTAGACCAAA 1020

AGAAGAAACG TGCTTGTCTC TGTATACCCG CAGAATGAAG TTACTGTTGT TAAAACCGGA 1080

TTTTTTCATT TTACTAGGTT CCGAAGAGTC CAGATGCTTG GTAGATGTTC AATACGTGAT 1140

TTTTTTTTTA ATTGAATGTG TTCATTTAAA ATCCTCCTTA ACATTTCTAG AAAGACTTCT 1200

TTCAATAAAT AATGGAATCT TAGAGGAAAA GTGGTTTTTT AAAAGCTAGG GAACTCCTCC 1260

ACTAAAAGTA ACCATTGGAA ACCTCGAATG AGGGCTAAAG TTTTAATCAT AAGAGAAAAG 1320

GCAGCATAAT GAAATGTGTA CACATACATA GTCAGTGGTC CATTTTAGGA AGCCAGTGGC 1380

GTCTGATAAA GAAATGTTAA GAGTAGTGAG GTTGAGGAAG GAAATTGTGG GGATTTGAAA 1440

TATTCTCTTT ATGTTGTTTC TCTTCTGAGT CATGGTAAAA CAATAAATTA TCATCTCTAG 1500

GTGGCAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1540

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 64 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met Lys Leu Leu Leu Leu Lys Pro Asp Phe Phe Ile Leu Leu Gly Ser
1 5 10 15

Glu Glu Ser Arg Cys Leu Val Asp Val Gln Tyr Val Ile Phe Phe Leu
20 25 30

Ile Glu Cys Val His Leu Lys Ser Ser Leu Thr Phe Leu Glu Arg Leu
35 40 45

Leu Ser Ile Asn Asn Gly Ile Leu Glu Glu Lys Trp Phe Phe Lys Ser
50 55 60

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ANTGACGCCTT TAGCTAGTCC TTCTATCA 29

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TNCAACAGTAT CAACCAGAAG TGCCAATC 29

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GNAAACAGTAT TAAATTGCAG AGTTCCAG 29

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CNAATCATCAT CTCGCTTACA CAGTCAGG 29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ANCGAGACAGA CCAAGGCTCA CAGGTGAA 29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GNGGACACACA CTCATAAAAC AGAAAATA 29

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

ANTAACCATAG AGTTTGGCTG GAGACCAC 29

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ANTCTTCCGAT GGCAACACCA GCCTGCTT 29

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GNTCACCATTT CTGTGTTCTT CAGTTTCT 29

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ANATTTAGGCT TATGCTGACC TCGTTGTA 29

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 162 amino acids

(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```
Met Gly Asn Ala Ser Tyr Ser Asp Ser Tyr Leu Glu Gly Ile Leu Leu
1               5                  10                  15

Lys Gly Val Phe Thr Cys Glu Pro Trp Ser Val Ser Val Gly Trp Ser
            20                  25                  30

Thr Ser Val Asp Cys Gly Trp Phe Gln Ser Ala Trp Leu Arg Ser Ser
        35                  40                  45

Thr Glu Gly Ser Thr Gly Ala Ala Leu Trp Gln Arg Leu Arg Arg Glu
    50                  55                  60

Met Asp Gln Pro Thr Pro Gly Thr Trp Leu His Gly Ile Arg Lys Gly
65                  70                  75                  80

Arg Cys Trp Pro Leu Cys Ser Cys Cys Leu Phe Leu Phe Leu Phe Ala
                85                  90                  95

Phe Asp Leu Val Ala Thr Asp Arg Val Ala Arg Asp Leu Val Phe Ser
            100                 105                 110

Ser Arg His Pro Ser Thr Pro Ala Leu Ser Gln Val Ser Cys Cys His
        115                 120                 125

Ala Asn Ala Met Ser Thr Leu Ala Pro Arg Pro Lys Ser Val Gln Arg
    130                 135                 140

Trp Pro Thr Xaa Ser Ser His Tyr Ser Leu Asn Ser Met Tyr His Leu
145                 150                 155                 160

Pro Leu
```

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 83 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```
Leu Val Gly Gly Pro Phe Leu Pro Pro Thr Ser Val Cys Cys Ser Cys
1               5                  10                  15

Gln Lys Thr Cys Leu Leu Leu Pro Ala Glu Arg Glu Gln Ser Pro Glu
            20                  25                  30

His His Gln Glu Cys Leu Leu Val Tyr Trp Gln Leu Ala Thr Pro Ser
        35                  40                  45

Ser Pro Phe Thr Gln Thr Arg Gly Arg Asp Gly Lys Gly Phe Phe Thr
    50                  55                  60

Glu His Ser Gly Thr Pro Tyr Arg Arg Lys Leu Asp Arg Leu Val Asn
65                  70                  75                  80

Gly Phe Ser
```

What is claimed is:

1. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 707 to nucleotide 1783;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 368 to nucleotide 838;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bp783_3 deposited under accession number ATCC 98369;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bp783_3 deposited under accession number ATCC 98369;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bp783_3 deposited under accession number ATCC 98369;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bp783_3 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 174 to amino acid 183 of SEQ ID NO:2;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

2. The polynucleotide of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.

3. A host cell transformed with the polynucleotide of claim 2.

4. The host cell of claim 3, wherein said cell is a mammalian cell.

5. A process for producing a protein encoded by the polynucleotide of claim 2, which process comprises:

(a) growing a culture of the host cell of claim 3 in a suitable culture medium; and

(b) purifying said protein from the culture.

6. A protein produced according to the process of claim 5.

7. The protein of claim 6 comprising a mature protein.

8. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 44;

(c) fragments of the amino acid sequence of SEQ ID NO:2 comprising the amino acid sequence from amino acid 174 to amino acid 183 of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone bp783_3 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins.

9. The protein of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.

10. The protein of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 44.

11. A composition comprising the protein of claim 8 and a pharmaceutically acceptable carrier.

12. A method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition of claim 11.

13. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1.

14. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 99 to nucleotide 1514;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 171 to nucleotide 1514;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 57 to nucleotide 623;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bu45_2 deposited under accession number ATCC 98369;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bu45_2 deposited under accession number ATCC 98369;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bu45_2 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bu45_2 deposited under accession number ATCC 98369;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 231 to amino acid 240 of SEQ ID NO:4;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

15. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 175;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising the amino acid sequence from amino acid 231 to amino acid 240 of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone bu45_2 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins.

16. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:3.

17. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 87 to nucleotide 980;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 147 to nucleotide 980;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ct864_4 deposited under accession number ATCC 98369;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ct864_4 deposited under accession number ATCC 98369;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ct864_4 deposited under accession number ATCC 98369;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ct864_4 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 144 to amino acid 153 of SEQ ID NO:6;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

18. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 189 to amino acid 290;

(c) fragments of the amino acid sequence of SEQ ID NO:6 comprising the amino acid sequence from amino acid 144 to amino acid 153 of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone ct864_4 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins.

19. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:5.

20. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 242 to nucleotide 580;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 1 to nucleotide 387;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone df396_1 deposited under accession number ATCC 98369;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone df396_1 deposited under accession number ATCC 98369;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone df396_1 deposited under accession number ATCC 98369;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone df396_1 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 51 to amino acid 60 of SEQ ID NO:8;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

21. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 48;

(c) fragments of the amino acid sequence of SEQ ID NO:8 comprising the amino acid sequence from amino acid 51 to amino acid 60 of SEQ ID NO:8; and

(d) the amino acid sequence encoded by the cDNA insert of clone df396_1 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins.

22. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:7.

23. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 236 to nucleotide 1213;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 1386 to nucleotide 1833;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dh1135_9 deposited under accession number ATCC 98369;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dh1135_9 deposited under accession number ATCC 98369;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dh1135_9 deposited under accession number ATCC 98369;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dh1135_9 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 157 to amino acid 166 of SEQ ID NO:10;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

24. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) the amino acid sequence of SEQ ID NO:31 from amino acid 1 to amino acid 147;

(c) fragments of the amino acid sequence of SEQ ID NO:10 comprising the amino acid sequence from amino acid 157 to amino acid 166 of SEQ ID NO:10; and

(d) the amino acid sequence encoded by the cDNA insert of clone dh1135_9 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins.

25. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:9.

26. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 334 to nucleotide 675;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 409 to nucleotide 675;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dn809_5 deposited under accession number ATCC 98369;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dn809_5 deposited under accession number ATCC 98369;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dn809_5 deposited under accession number ATCC 98369;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dn809_5 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising the amino acid sequence from amino acid 52 to amino acid 61 of SEQ ID NO:12;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

27. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:12;

(b) the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 110;

(c) fragments of the amino acid sequence of SEQ ID NO:12 comprising the amino acid sequence from amino acid 52 to amino acid 61 of SEQ ID NO:12; and

(d) the amino acid sequence encoded by the cDNA insert of clone dn809_5 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins.

28. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:11.

29. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 447 to nucleotide 791;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 597 to nucleotide 791;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1 to nucleotide 546;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ej224_1 deposited under accession number ATCC 98369;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ej224_1 deposited under accession number ATCC 98369;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ej224_1 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ej224_1 deposited under accession number ATCC 98369;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising the amino acid sequence from amino acid 52 to amino acid 61 of SEQ ID NO:14;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

30. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:14;

(b) the amino acid sequence of SEQ ID NO:14 from amino acid 82 to amino acid 100;

(c) fragments of the amino acid sequence of SEQ ID NO:14 comprising the amino acid sequence from amino acid 52 to amino acid 61 of SEQ ID NO:14; and

(d) the amino acid sequence encoded by the cDNA insert of clone ej224_1 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins.

31. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:13.

32. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 18 to nucleotide 347;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 1 to nucleotide 345;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ek591_1 deposited under accession number ATCC 98369;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ek591_1 deposited under accession number ATCC 98369;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ek591_1 deposited under accession number ATCC 98369;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ek591_1 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising the amino acid sequence from amino acid 50 to amino acid 59 of SEQ ID NO:16;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

33. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 109;

(c) fragments of the amino acid sequence of SEQ ID NO:16 comprising the amino acid sequence from amino acid 50 to amino acid 59 of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone ek591_1 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins.

34. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:15.

35. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 593 to nucleotide 1663;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 833 to nucleotide 1663;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 648 to nucleotide 1063;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone er381_1 deposited under accession number ATCC 98369;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone er381_1 deposited under accession number ATCC 98369;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone er381_1 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone er381_1 deposited under accession number ATCC 98369;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising the amino acid sequence from amino acid 173 to amino acid 182 of SEQ ID NO:18;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

36. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:18;

(b) the amino acid sequence of SEQ ID NO:18 from amino acid 20 to amino acid 157;

(c) fragments of the amino acid sequence of SEQ ID NO:18 comprising the amino acid sequence from amino acid 173 to amino acid 182 of SEQ ID NO:18; and

(d) the amino acid sequence encoded by the cDNA insert of clone er381_1 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins.

37. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:17.

38. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 1055 to nucleotide 1246;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 759 to nucleotide 1152;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone gq38_1 deposited under accession number ATCC 98369;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone gq38_1 deposited under accession number ATCC 98369;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone gq38_1 deposited under accession number ATCC 98369;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone gq38_1 deposited under accession number ATCC 98369;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment comprising the amino acid sequence from amino acid 20 to amino acid 29 of SEQ ID NO:20;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

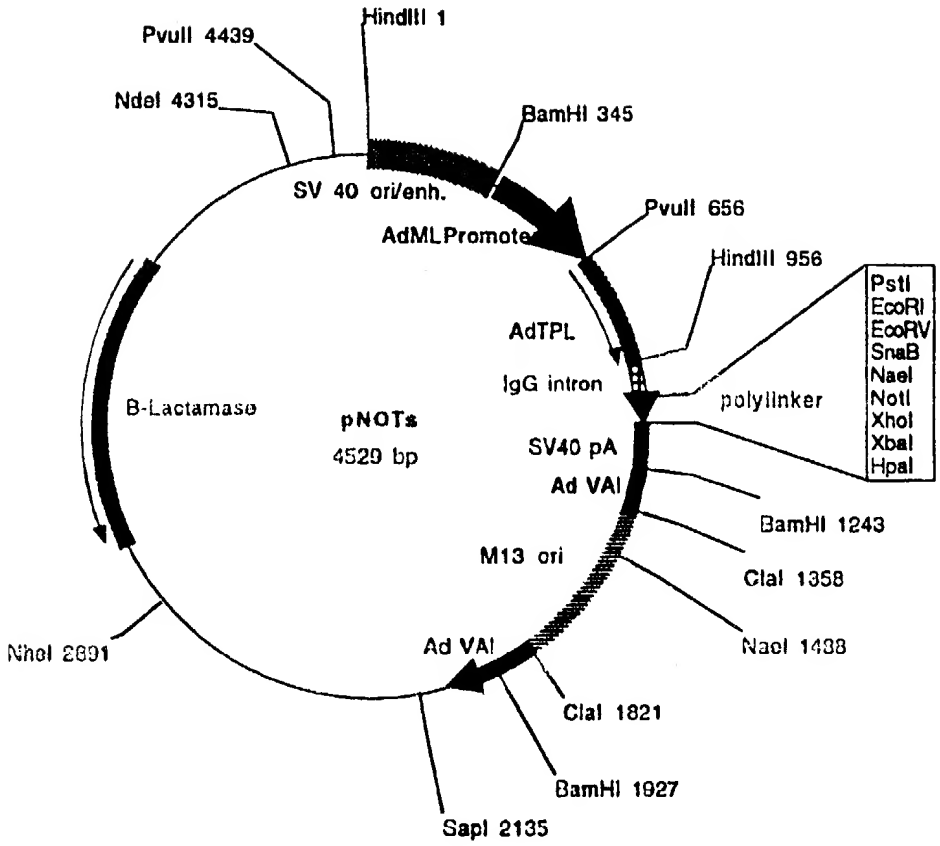
(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

39. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:20;

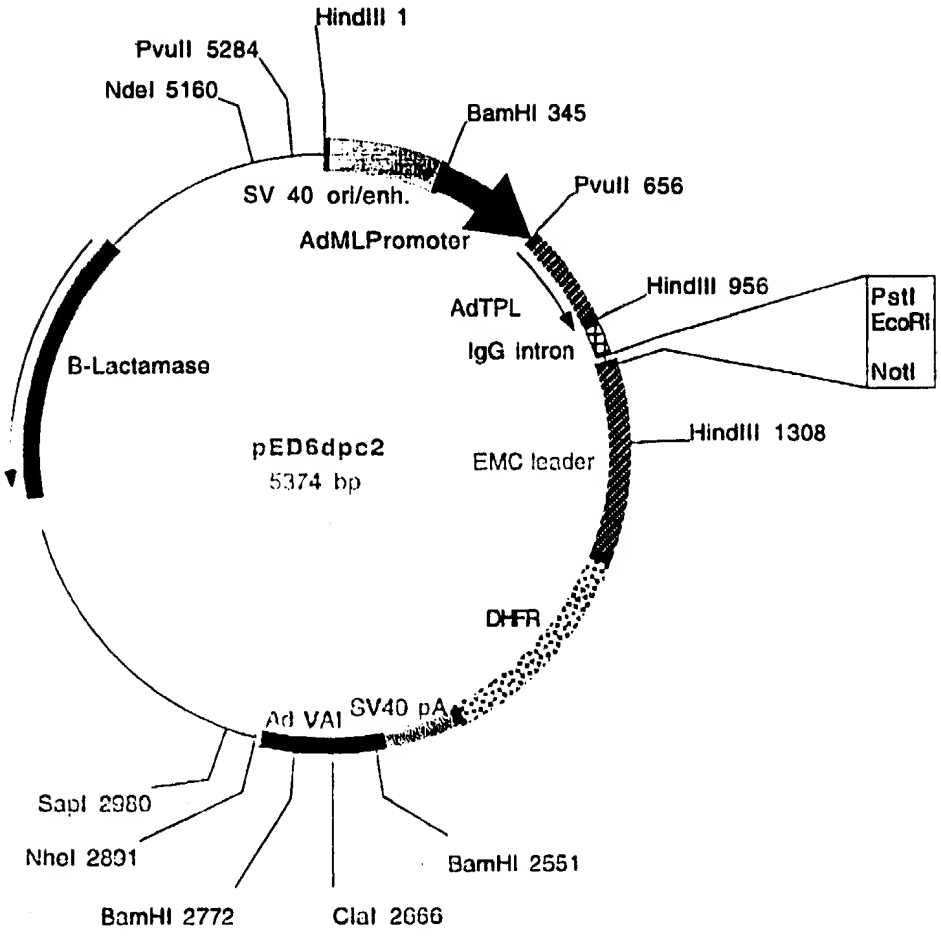
FIGURE 1B



Plasmid name: pNOTs
Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al,1989. Mol.Cell.Biol.9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the ClaI site. SST cDNAs are cloned between EcoRI and NotI

FIGURE 1A



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

(b) the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 32;

(c) fragments of the amino acid sequence of SEQ ID NO:20 comprising the amino acid sequence from amino acid 20 to amino acid 29 of SEQ ID NO:20; and

(d) the amino acid sequence encoded by the cDNA insert of clone gq38_1 deposited under accession number ATCC 98369;

the protein being substantially free from other mammalian proteins.

40. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:19.